

PUBLICACIONES DEL INSTITUTO CARO Y CUERVO

SERIES MINOR

XLIX

---

JOSÉ ALEJANDRO CORREA DUARTE

MANUAL DE  
ANÁLISIS ACÚSTICO  
DEL HABLA CON PRAAT



BOGOTÁ  
2014

# MANUAL DE ANÁLISIS ACÚSTICO DEL HABLA CON PRAAT



PUBLICACIONES DEL INSTITUTO CARO Y CUERVO  
SERIES MINOR  
XLIX

---

JOSÉ ALEJANDRO CORREA DUARTE

# MANUAL DE ANÁLISIS ACÚSTICO DEL HABLA CON PRAAT



BOGOTÁ  
2014



Título: – Manual de análisis acústico del habla con Praat. – Autor: José Alejandro Correa Duarte. – Bogotá: Instituto Caro y Cuervo. Imprenta Patriótica, 2014. 130 pp. (Series Minor; 49)

Contenido: Fonética acústica. – Praat: doing Phonetics by Computer. – Fonética experimental. – Lingüística de corpus. – Informática aplicada a la investigación lingüística.

Instituto Caro y Cuervo

Directora General: Carmen Millán de Benavides

Dirección Editorial: César Buitrago

Primera edición 2014

Diseño de carátula: Víctor Galvis R.

Diagramado por el autor con L<sup>A</sup>T<sub>E</sub>X

Corrección: Alejandro Molina

Carátula: Imprenta Patriótica

Serie: Minor XLIX

ISBN: En trámite

Hecho en Colombia

Enlace:



<http://www.bibliodigitalcaroycuervo.gov.co/998/>

Biblioteca Digital Palabra, Instituto Caro y Cuervo

Licencia:



<http://creativecommons.org/licenses/by-nc/4.0/>

## INTRODUCCIÓN

Praat es un programa de libre acceso creado por Paul Boersma y David Weenink (2014) para analizar, sintetizar y manipular señales de habla. El programa se puede recuperar en «<http://www.praat.org/>» para los sistemas operativos Linux, Macintosh y Windows. Las principales funciones de Praat son:

1. ANÁLISIS DEL HABLA: el programa ofrece opciones para procesar muestras de habla y obtener información acústica como espectrogramas de banda estrecha y banda ancha, espectros, formantes, frecuencia fundamental ( $F_0$ ), intensidad, duración, entre otros. Esta función sirve para estudiar la estructura fonética de una lengua, fenómenos dialectales o estudiar los rasgos fonéticos idiosincrásicos de un individuo.
2. SÍNTESIS DEL HABLA: el objeto **KlattGrid** (Weenink, 2009) es un sintetizador por formantes inspirado en el modelo de Klatt y Klatt (1990). Con Praat también es posible realizar síntesis articulatoria según la fonología funcional de Boersma (1998), resíntesis de  $F_0$  para estudiar la entonación y convertir texto a voz con el sintetizador *e-speak*.
3. TRANSCRIPCIÓN: con un objeto **TextGrid** se transcriben y etiquetan archivos de audio para crear corpus orales con diferentes grados de detalle y ajustados a diferentes propósitos como la documentación de lenguas y la experimentación.

4. MANIPULACIÓN DE SEÑALES: con la función **To manipulation** se modifica la duración, la curva de  $F_0$  y la intensidad de los sonidos del habla. Además, Praat tiene opciones de filtrado y normalización de señales.
5. SCRIPTS: con el lenguaje de programación de Praat se automatizan tareas y se evita la repetición manual de procedimientos. Es decir, las mediciones realizadas sobre grandes bases de datos no requieren el uso reiterado de las mismas opciones del menú (Wempe y Boersma, 2003). Otra ventaja de los scripts es que podemos utilizar códigos que ponen a disposición diferentes investigadores y descubrir trucos para simplificar el número de líneas de las rutinas creadas por nosotros mismos.
6. FIGURAS: con la ventana de imágenes o **Praat Picture** se rotulan y decoran figuras para incluir en tesis, artículos, libros o presentaciones. La herramienta es flexible y varía fácilmente las dimensiones, la tipografía y los colores.
7. ESTADÍSTICA: es posible aplicar métodos estadísticos (*análisis de escalamiento multidimensional*, *análisis de componentes principales* y *análisis discriminante*) a tablas con datos acústicos.
8. EXPERIMENTOS: la herramienta **PraatMFC** (*Multiple forced choice listening experiment*) sirve para diseñar experimentos. Esta función es ideal para quienes quieren estudiar la percepción de los sonidos del habla y necesitan crear estímulos, organizarlos, presentarlos al participante y almacenar las respuestas.
9. ALGORITMOS DE APRENDIZAJE: permite usar redes neurales (*feedforward neural networks*) y simular gramáticas (ordinales, estocásticas, armónicas) según la teoría de la optimalidad.

La documentación oficial de Praat (Boersma y Weenink, 2014) explica en detalle estas funciones y la manera como el programa procesa una señal de audio. En Internet también hay un buen número de documentos —principalmente escritos en inglés— para aprender los fundamentos del programa, y hay un grupo de discusión que permite a los usuarios formular dudas, ofrecer soluciones, proponer mejoras y compartir información. Aconsejo a los lectores inscribirse en la lista de miembros para estar al tanto de los problemas frecuentes y su posible resolución «<http://uk.groups.yahoo.com/group/praat-users/>».

La razón que me llevó a escribir este manual de Praat fue el deseo de contar con una herramienta apropiada para impartir cursos de fonética experimental y al mismo tiempo estimular el autoaprendizaje de la disciplina. Para llevar a cabo esta labor, estudié la documentación de Praat (Boersma y Weenink, 2014), ponderé la información, la confronté con mi experiencia como usuario y, finalmente, trasladé sus líneas esenciales a un lenguaje menos esotérico.

En estas páginas el lector encontrará los procedimientos básicos para analizar, transcribir y procesar señales con Praat. La selección de contenidos está diseñada para que el estudiante desarrolle habilidades básicas en el análisis del habla y comprenda a cabalidad los fundamentos de la fonética acústica. Por esta razón, se hace énfasis en la medición manual de los parámetros y no en la elaboración de scripts o en la obtención de resultados rápidos. En el capítulo 1 se explican las operaciones básicas del programa tales como abrir, guardar archivos, grabar sonidos de audio, ejecutar un script y dibujar ondas; en el capítulo 2 se indica cómo ajustar los parámetros para realizar análisis espectrográficos (espectrogramas y espectros), ajustar los contornos de  $F_0$ , la intensidad y los pulsos glotales; en el capítulo 3 se aborda el tema de la transcripción de los fenómenos segmentales y supra-segmentales y, finalmente, en el capítulo 4 se presentan algunas técnicas de procesamiento y manipulación digital del habla.

Sugiero usar este manual para poner en práctica los conceptos estudiados en clase o en obras introductorias a la fonética; al final de cada capítulo el lector encontrará una sección con fuentes bibliográficas que podrá usar para este fin. Además, en algunos capítulos se hace referencia a otros manuales que tratan temas complementarios. Las personas interesadas en aprender funciones avanzadas o los fundamentos matemáticos de Praat deben consultar la página oficial del programa o el libro en curso de David Weenink (2014).

Este trabajo es producto de las actividades académicas del autor dentro del grupo de *Lingüística de Corpus* del Instituto Caro y Cuervo. Parte de los capítulos surgieron de la interacción con varios grupos de investigación como *Bilingüismo y lenguas en contacto*, *Lenguaje, encuentros y diversidad*, y *Aslec*, luego fueron implementados en cursos cortos de fonética en el Diplomado en Lingüística del Instituto Caro y Cuervo, en la Maestría en Lingüística de la Universidad Nacional de Colombia, y actualmente se usa en la Maestría en Lingüística del Seminario Andrés Bello del Instituto Caro y Cuervo. Agradezco a mis colegas y alumnos por darle sentido a esta idea –espero que estas páginas ayuden al desarrollo de sus labores académicas– y a Linda Rodríguez y Eduardo Dominiccini por los valiosos comentarios. Finalmente, quiero expresar mi gratitud con la Dirección General del Instituto Caro y Cuervo por su apoyo editorial y a mis maestros Jorge Rico Ródenas, Juana Gil Fernández y Joaquim Llisterri del Programa Oficial de Posgrado en Estudios Fónicos (CSIC/UIMP) por las enseñanzas fundamentales.

JOSÉ ALEJANDRO CORREA DUARTE  
Seminario Andrés Bello  
Instituto Caro y Cuervo

## CONVENCIONES

**Mecanográfico**

Se usa para indicar opciones del menú, tipos de objeto y atajos.

*Cursivas*

Conceptos clave. p. ej., *frecuencia de muestreo*.

**Enlace**

En la versión electrónica del manual se utiliza para citar secciones de la documentación oficial de Praat (Boersma y Weenink, 2014) p. ej., [Intro 5. Formant analysis](#). Para acceder al manual oficial remítase a la página de Internet o seleccione **Help → Go to manual page...** en la ventana de objetos.

→

Como es costumbre, la flecha significa ‘elegir en el (sub)-menú la opción *x*’. Así, **New → Record mono Sound...** quiere decir: elegir **New** y luego **Record mono Sound...** en el menú.



# 1 FUNDAMENTOS

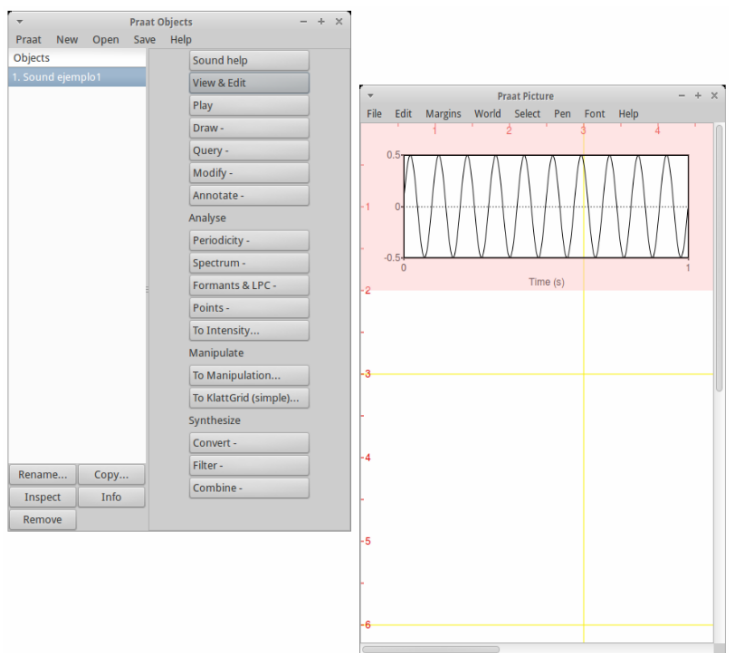
## 1.1 INTERFAZ GRÁFICA

Al abrir Praat aparece la ventana de objetos (**Praat Objects**) y la ventana de imágenes (**Praat Picture**), como se ve en la figura 1.1. En la parte superior de cada ventana hay un *menú estático* con opciones para crear, abrir y guardar archivos de audio y texto. El entorno visual o interfaz gráfica de usuario (GNU) depende del sistema operativo que esté instalado en el equipo: los pantallazos presentados en estas páginas, por ejemplo, se capturaron en un computador con sistema Linux.

La ventana **Praat Objects** está conformada por dos columnas. En la primera se listan los objetos en uso y en la segunda hay un *menú dinámico* que sirve para analizar (**Analyse**), manipular (**Manipulate**) o sintetizar (**Sinthesize**) sonidos. Este menú se puede modificar, así que es posible añadir botones para realizar tareas ajustadas a nuestras necesidades. En la parte inferior de **Praat Objects** hay botones para buscar (**Inspect**) u obtener información (**Info**), copiar (**Copy...**), cambiar el nombre de un objeto (**Rename...**) o borrarlo (**Remove**). En la parte superior derecha del menú el usuario encontrará la opción **Help**, la cual le permitirá acceder a la sección de preguntas frecuentes o realizar búsquedas en el manual.

La ventana **Praat Picture** tiene una cuadrícula para diseñar las figuras de nuestros artículos, tesis o libros. Desde el menú estático que está ubicado en la parte superior de esta ventana se modifican elementos visuales como el texto, se añaden etiquetas fonéticas, se cambia el tipo y ancho de línea, el tipo y tamaño de fuente, los colores, se introducen marcas temporales o nominales en la abscisa y la ordenada, etc.





**Figura 1.1.** Ventana de objetos (izquierda) y ventana de imágenes (derecha)

## 1.2 ABRIR Y GUARDAR ARCHIVOS

En este apartado veremos cómo abrir datos almacenados en nuestro equipo y cómo guardar cambios sobre archivos existentes. Praat lee los formatos de audio que usan los principales sistemas operativos (AIFF, AIFC, WAV, AU, FLAC y decodifica MP3). Para leer archivos de texto y archivos de audio de corta duración simplemente debemos elegir:

Praat Objects → Read → Read from file...

Cuando se realiza el análisis acústico de una lengua o dialecto (p. ej., para conocer su inventario de sonidos) se trabaja con enunciados cortos que contienen información controlada. No obstante, es común usar archivos extensos cuando se estudia habla espontánea, conversaciones o una muestra de tradición oral. Para abrir archivos de audio largos pero con un tamaño menor a 2Gb elija:

**Praat Objects → Read → Open long sound file...**

Todos los archivos que se abren durante una sesión de trabajo se listan en orden de aparición en la ventana de objetos (**Praat Objects**). Debe tenerse en cuenta que el programa no guarda por defecto los cambios realizados. Para guardar o escribir archivos de texto con datos cuantitativos o información de una señal escoja:

**Praat Objects → Save as text file...**

Para guardar un archivo de sonido debe elegirse el tipo de formato (depende del sistema operativo del usuario) y la ubicación de la carpeta en la que se almacenarán los datos.

**Praat Objects → Save as WAV file... etc.**

No sobra insistir en que es necesario guardar los objetos nuevos o cada vez que se realicen cambios relevantes a objetos existentes. Si bien perder horas de trabajo es bastante instructivo, puede tener consecuencias negativas en una sociedad que favorece la competitividad y sanciona el uso intangible del tiempo.

### 1.3 GRABACIÓN

Para grabar en el computador deben configurarse las propiedades de sonido y conectar un micrófono al equipo. Si desea que las señales capturadas tengan la calidad suficiente para realizar un

análisis acústico, es necesario usar una tarjeta de sonido externa o una interfaz de audio para controlar la señal de entrada y evitar la interferencia de los componentes del equipo.

Si desea grabar con Praat vaya a la ventana de objetos y elija **New** → **Record mono Sound...** o simplemente (**Ctrl + R**). Luego aparecerá una ventana en la que debe ajustar los siguientes parámetros:

1. **Channels: Mono.**
2. **Sampling frequency:** para grabar voz humana la frecuencia de muestreo recomendable es 44.100Hz (opción por defecto en Praat).
3. **Name:** se asigna un nombre que luego facilite la búsqueda del archivo.
4. **Record:** inicia la grabación.
5. **Meter:** para lograr una grabación con los niveles de intensidad adecuados, la señal no puede superar el área verde; las variaciones entre el área amarilla y roja indican que debe reducirse el nivel de intensidad.
6. **Stop:** detiene la grabación.
7. **Play:** escuche la muestra y repítala si es necesario.
8. **Save to list & close:** la muestra grabada aparece en la primera columna de la ventana de objetos. Si se elige **Close**, el panel de grabación se cierra pero no se guarda el archivo.

Con estos ajustes Praat graba 220 segundos. Para grabar muestras de mayor longitud es necesario ir a **Praat** → **Preferences** → **Sound recording preferences...** → **Buffer size (MB)** y asignar un valor superior a 20 MB (véase [Intro 1.1. Recording a sound](#)).

## 1.4 EL OBJETO SOUND

Praat es un programa orientado a objetos, lo cual significa que antes de realizar cualquier tarea se debe seleccionar el objeto i. e., *la selección precede a la acción* (Weenink, 2014, p.7). Existen diferentes tipos, dentro de los más usados cabe mencionar **Sound**, **LongSound**, **Table**, **To manipulation**, **PointProcess** y **TextGrid**. Esta sección explica cómo extraer información del objeto **Sound** (i. e., un archivo de audio).

Cuando se abre un objeto **Sound** se activa el menú dinámico de **Praat Objects** y se enumera de acuerdo con su orden de aparición. Para saber los datos básicos de la señal tan solo debe oprimir **Info** en la parte inferior de la ventana y aparecerá un texto (**Praat Info**) con los datos que se muestran en la ventana 1.1. En las líneas 1-4 se presentan datos generales como el código del objeto, su tipo, nombre y la fecha. En la línea 5 se muestra el número de canales; si el archivo está grabado en mono aparece el número 1 y si está grabado en estéreo aparecerá el número 2. Luego el cuadro muestra el tiempo de inicio del sonido, el tiempo de finalización y la duración total (líneas 6-9).

Entre las líneas 10-14 el usuario encontrará el número de muestras (*sampling*) utilizadas para procesar digitalmente la señal, el periodo de las muestras, la tasa de muestreo y el punto en que se tomó el primer valor (recuerde que para grabar un archivo de audio es recomendable utilizar una tasa de 44.100 Hz). La amplitud (líneas 15-19) se expresa en pascales (Pa) y en el cuadro podremos obtener su rango (valor máximo y mínimo), valor medio y media cuadrática (*Root mean square*) (si su sonido tiene una amplitud inferior a 1.0Pa es probable que no sea completamente audible. Para modificar la amplitud debe elegir **Modify** → **Scale peak** y ajustar el valor a 0.99). Las líneas finales del cuadro muestran otros datos como son la energía total, la media de la potencia y la desviación estándar del canal.

**Ventana 1.1. Praat Info**

```

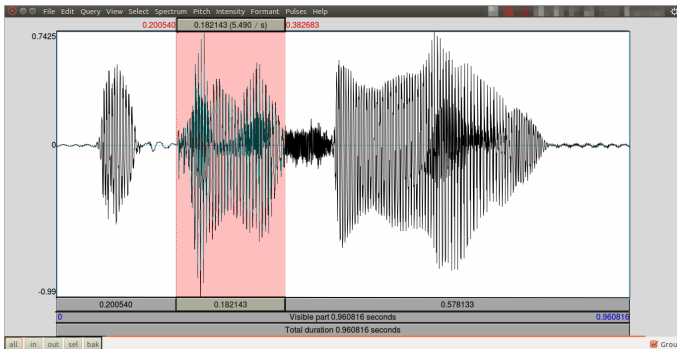
1 Object id:
2 Object type:
3 Object name:
4 Date:
5 Number of channels:
6 Time domain:
7     Start time:
8     End time:
9     Total duration:
10 Time sampling:
11     Number of samples:
12     Sampling period:
13     Sampling frequency:
14     First sample centred at:
15 Amplitude:
16     Minimum:
17     Maximum:
18     Mean:
19     Root-mean-square:
20 Total energy:
21 Mean power (intensity) in air:
22 Standard deviation in channel:

```

Si desea obtener estos valores individualmente vaya al menú y despliegue la opción **Query**. Cada una de las opciones del menú remite a los campos que se muestran entre las líneas 1-22 del cuadro **Praat Info**. Las opciones restantes **Get nearest zero crossing...** y **Get absolute extremum...** hacen referencia al punto del tiempo en que la amplitud de la onda no es positivo ni negativo y al valor máximo (expresado en Pa) en una ventana o en la totalidad del sonido.

## 1.5 EDITOR DE SONIDO

Para abrir una señal de habla en el Editor se selecciona el objeto **Sound** y luego el botón **View & Edit**. En la ventana emergente se ve la forma de onda (figura 1.2) y es posible seleccionar parte de un sonido, escucharlo y obtener su duración. En el ejemplo se realizó una selección de 0.18s (área resaltada) en una muestra con una duración total de 0.96s. En la parte superior de la forma de onda se indica el inicio de la selección (0.20s), la duración total, su inversa (0.18 (5.490/s)) y el tiempo de corte (0.38s). En la parte inferior del Editor hay tres barras que permiten escuchar la selección, la señal visible en la ventana (**Visible part**) y la totalidad del archivo (**Total duration**).



**Figura 1.2.** Editor de sonido

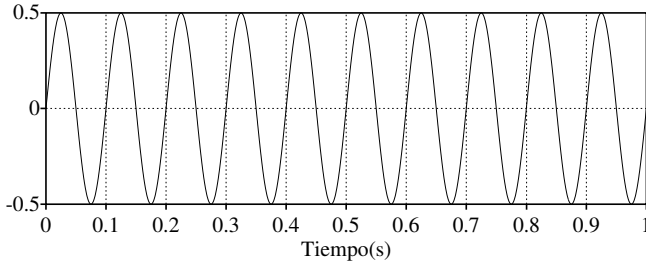
En la parte inferior izquierda, los botones **all**, **in**, **out**, **sel** y **bak** sirven para aumentar o alejar el detalle de la imagen. Adicionalmente, en el costado inferior derecho la opción **Group** asegura que, si tenemos dos versiones del mismo objeto, las selecciones se realizan simultáneamente en ambas ventanas.

En la parte superior del Editor hay un menú que ofrece las siguientes opciones de edición y búsqueda:

1. **File:** es posible dibujar la forma de onda que se muestra en el Editor o una selección determinada usando las opciones **Draw visible sound...** y **Draw selected sound...**. Para cortar un archivo largo en pequeños trozos puede usar **Extract selected sound (preserve times)** o **Extract selected sound (time from 0)**.
2. **Edit:** en este menú hay una serie de opciones para cortar (**cut**), copiar (**Copy selection to Sound clipboard**) y pegar (**Paste after selection**) durante la edición de muestras y la creación de estímulos. Con **Set selection to zero** se anula la intensidad de la señal y con **Reverse selection** se invierte la onda seleccionada.
3. **Query:** la opción **Editor info** da información detallada sobre los parámetros del Editor de sonido, en tanto, **Sound info** genera la información que está en la ventana 1.1. Hay otras alternativas como **Get start selection**, **Get end of selection** y **Get selection length** para conocer el inicio, el final y la duración total de un área seleccionada. También es posible crear registros (**Log files**) en un archivo de texto con mediciones acústicas (p. ej., *formantes*, *frecuencia*, *intensidad*).
4. **View:** esta parte del menú ofrece diferentes alternativas para aumentar el detalle de la imagen i. e., **zoom**, **show all**, **zoom in**, **zoom out**, etc. Además, con **Play**, **Play or stop** e **Interrupt playing** se pueden escuchar fracciones específicas de la muestra analizada.
5. **Select:** sirve para mover el cursor al inicio o final de una selección, mover el cursor de izquierda a derecha o hacia el cruce por cero más próximo.

Si continuamos de izquierda a derecha en el menú del Editor, nos encontraremos con las opciones **Spectrum**, **Pitch**, **Intensity**, **Formants** y **Pulses**. Cada una de estas alternativas remite a un nivel de análisis acústico independiente. En el capítulo siguiente se hablará sobre la información lingüística que transmite cada uno, se explicará la manera de configurar los parámetros de análisis y la manera de acceder a la representación acústica desde el editor del programa.

## 1.6 CÓMO CREAR ONDAS SINUSOIDALES



**Figura 1.3.** Onda sinusoidal de 10Hz

Con la aplicación de una fórmula se crean ondas sinusoidales con una *frecuencia* y *amplitud* determinadas. Recordemos que la frecuencia se define como el número de veces que se repite un ciclo en una unidad de tiempo (segundo). La onda de la figura 1.3, por ejemplo, tiene una frecuencia de 10Hz porque realiza diez ciclos de 0,1s en un segundo. Esto se puede saber calculando la relación entre frecuencia (F) y periodo (T). Si  $F = \frac{1}{T}$ , entonces  $\frac{1}{0,1} = 10\text{Hz}$ . De otro lado, la amplitud es el rango de variación de la presión del aire y, como vimos en páginas anteriores, se mide en Pa (Ladefoged, 1996, p.16). Como se ve en la figura 1.3 la amplitud varía entre 0,5Pa y -0,5Pa.



Para crear una onda sinusoidal vaya a **Praat Objects → New → Sound → Create Sound from formula...** En este punto aparecerá un formulario con los parámetros necesarios para crear el sonido. Por ejemplo, para hacer una onda de 10 hercios (Hz), con una presión sonora entre +0.5Pa y -0.5Pa y una duración de 1s, debemos modificar la configuración así:

1. **Name:** se asigna un nombre que facilite la identificación de la información.
2. **Channels:** Mono.
3. **Start time (s):** tiempo de inicio 0.
4. **End time (s):** tiempo de finalización 1.
5. **Sampling frequency:** 44100 Hz.
6. **Formula:**  $1/2 * \sin(2\pi * 10 * x)$ .
7. **OK:** se realiza la tarea y se lista el objeto en la ventana de objetos.

En la fórmula,  $1/2$  es la *amplitud*,  $2\pi$  es el *periodo* de la función sinusoidal, 10 es la *frecuencia* y  $x$  representa tramos temporales entre 0 y 1s (véase Weenink, 2014, p.15-18). Si el lector desea crear una onda con la frecuencia más alta, tan solo tiene que cambiar el valor de esta *variable* en la fórmula.

Una vez hecho esto podemos automatizar el procedimiento con un script i. e., un conjunto de líneas de código que permiten realizar una tarea sin necesidad de usar los botones del menú. Para ello se debe ir a **Praat Objects → praat → New Praat script**. En ese punto aparecerá una ventana de texto; en el menú de la parte superior elija **Edit** y después **Paste History (Ctrl + H)**. Como consecuencia, obtendrá un resumen de las acciones descritas en 1-7. Es probable que haya más información de la necesaria, así que debemos estar seguros de dejar tan solo

los datos relevantes como se ve en el **Script 1a** (la numeración no hace parte del código). Para ejecutar las órdenes oprima **Run** o **Ctrl + R**. Para guardar el script vaya a **Script → File → Save as...**

```
1 Create Sound from formula: "ejemplo1", 1, 0, 1,  
2 ... 44100, "1/2 * sin(2*pi*10*x)"
```

### Ventana 1.2. Script 1a

Estas líneas pueden mejorarse, de tal suerte que sea posible crear ondas que se diferencien por su *frecuencia* o algún otro parámetro como la *amplitud*. Lo primero que se hace es asignar un título (lo debe preceder el signo **#** para indicar que es un comentario) y, segundo, definir la variable que desea controlar. En el **Script 1b** la variable frecuencia (*f*) se define en la línea 2 y luego se introduce en la fórmula como se aprecia al final de la línea 4. Finalmente, se añade la orden **Play** para reproducir el objeto una vez ha sido creado por Praat.

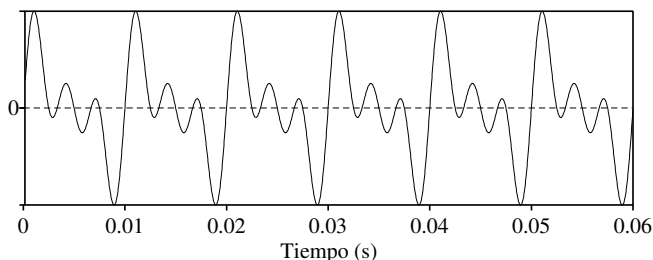
```
1 #onda sinusoidal  
2 f=10  
3 Create Sound from formula: "ejemplo1", 1, 0, 1,  
4 ... 44100, "1/2 * sin(2*pi*f*x)"  
5 Play
```

### Ventana 1.3. Script 1b

## 1.7 CÓMO CREAR ONDAS COMPLEJAS

El teorema de Fourier afirma que las *ondas complejas* son la combinación de dos o más tonos puros i. e., la sumatoria de las presiones positivas y la sustracción de las presiones negativas

de cada intervalo de tiempo (Ladefoged, 1996, p.36). Las ondas complejas tienen una *frecuencia fundamental* ( $F_0$ ) y un conjunto de *armónicos* que son múltiplos enteros de esta. Este concepto es importante en fonética porque ha sido aplicado con éxito al análisis de sonidos del habla como las vocales.



**Figura 1.4.** Onda compleja con componentes de 100Hz, 200Hz, 300Hz

Para crear una onda compleja con  $F_0$  de 100Hz y armónicos de 200Hz y 300Hz (figura 1.4) es necesario ir a **Praat Objects** → **New** → **Sound** → **Create Sound from tone complex...** y ajustar los siguientes parámetros:

1. **Name:** se asigna un nombre que facilite la búsqueda de la información.
2. **Start time (s):** tiempo de inicio 0.
3. **End time (s):** tiempo de finalización 1.
4. **Sampling frequency (Hz):** 44.100Hz.
5. **Phase:** especificar si se quiere utilizar la función seno o coseno. Como indican los autores del programa, esta diferencia no provoca diferencias perceptivas.

6. **Frequency step (Hz)**: la distancia entre los componentes es igual a la  $F_0$ .
7. **First frequency (Hz)**:  $0 = \text{Frequency step(Hz)}$ .
8. **Ceiling (Hz)**: es la frecuencia máxima considerada en el análisis. Si el valor asignado es 0, se descartarán las frecuencias superiores a la *frecuencia Nyquist*. Por su parte, la frecuencia Nyquist es la mitad de la tasa de muestreo i.e., 22.050Hz.
9. **Number of components**: el número de ondas que queremos en nuestra onda= 3.
10. **Ok**: se realiza la tarea y se lista el objeto en la ventana de objetos.

Como se mostró en la sección anterior, es posible reducir un conjunto de acciones a unas líneas de código. Los parámetros descritos en 1-10 se simplifican con el **Script 2a**:

```
1 #Onda compleja f0 = 100Hz
2 f=100
3 Create Sound from tone complex: "ejemplo_2", 0, 1,
4 ... 44100, "Sine", f, 0, 0, 3
5 Play
```

#### Ventana 1.4. Script 2a

En la línea 1 se define el nombre del script y en la línea 2 se define la variable *frecuencia fundamental* ( $f=100$ ). En la línea 3 se crea el sonido complejo, se asigna el nombre del objeto (ejemplo2) y la duración entre 0 y 1 segundo. La línea 4 inicia con tres puntos (...) indicando que la línea de código es continuación de la anterior. Si desea integrar la línea 4 con la línea tres elimine

los tres puntos (...) y asegúrese de diferenciar cada dato con una coma (.). Luego se elige la frecuencia de muestreo, la fase (**Sine**), se asigna a la primera frecuencia un valor de 0 ( $0 = f$ ), un techo (**Ceiling**) de 0 ( $0 = \text{frecuencia Nyquist}$ ) y 3 componentes de onda múltiplos de  $f$ . Finalmente, la última línea del script ordena reproducir (**Play**) el objeto creado.

El lector puede experimentar cambiando el valor de la variable  $f$ , el número de componentes de la onda (**Number of components**), la fase (**Phase**), etc. Si toma nota de los cambios acústicos observados, podrá aprender a usar los parámetros y comprender sus consecuencias a nivel auditivo.

La tarea de comparar los efectos perceptivos de las ondas se facilitaría si pudieramos repetir automáticamente la tarea descrita en el **script 2a**. De esta manera no sería necesario cambiar el valor de la frecuencia varias veces para hacer una comparación. Por ejemplo, podemos crear 3 ondas complejas con  $F_0$  de 100Hz, 200Hz y 300Hz y 3 armónicos (múltiplos enteros de cada valor de  $F_0$ ). El problema es hacer que el script ejecute tres veces la orden y en cada caso cambie el valor de la variable.

```

1  #Ondas complejas  f0 =100Hz, 200Hz y 300Hz.
2  for i from 1 to 3
3      f=i * 100
4      Create Sound from tone complex:
5      ... "ejemplo_3", 0, 1, 44100, "Sine", f,
6          0, 0, 3
7      Play
7  endfor

```

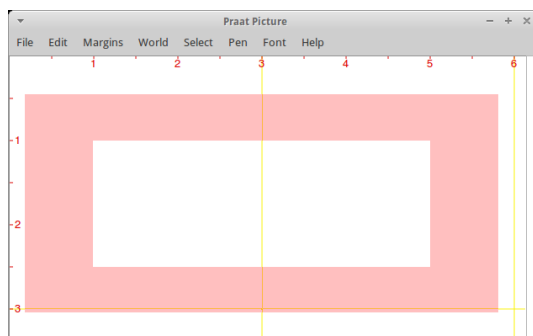
**Ventana 1.5. Script 2b**

Para resolver este problema se debe ajustar la sintaxis del script con un bucle (*loop*). Se trata de un procedimiento de programación en el cual la variable toma los valores que se encuentran entre dos expresiones numéricas. Por ejemplo, en la línea 2 del **script 2b** se indica que la variable *i* tomará los valores que se encuentran entre la expresiones 1-3. En la línea 3 se inicia el bucle con **for** y se redefine la variable *i*. e., se multiplica por 100 para que genere las frecuencias deseadas al realizar cada ciclo. En las líneas 4-6 se hacen las declaraciones necesarias para crear las ondas, y en la línea 7 se finaliza el bucle con **endfor**.

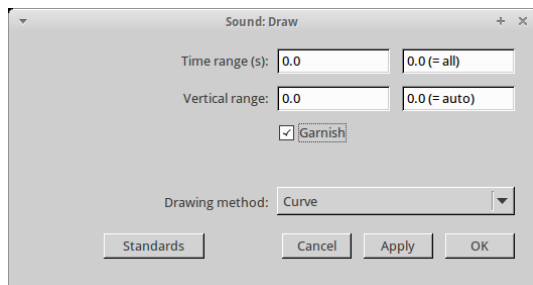
### 1.8 DIBUJANDO UNA ONDA

Una de las mejores funciones de Praat es la de crear imágenes de alta calidad. En este apartado describiré en tres pasos cómo dibujar una onda y rotularla. Para ilustrar este punto utilizaré la onda de 10 Hz que se muestra en la figura 1.3:

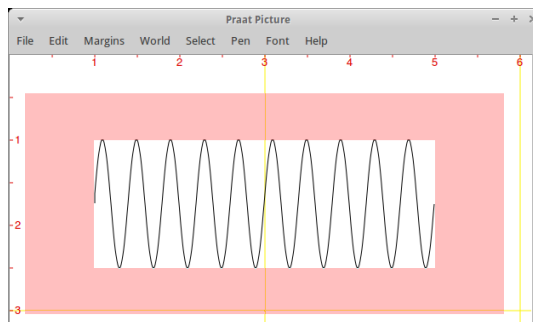
1. El primer paso es ir a la ventana de imágenes (**Praat Picture**) y seleccionar el área donde desea ubicar la figura. Para dibujar una onda o *forma de onda* es conveniente hacer una selección a manera de rectángulo.



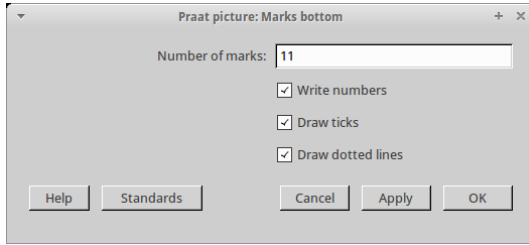
2. Segundo, seleccionar el sonido en **Praat Objects** y oprimir **Draw- → Draw ...** en el menú dinámico. En ese punto emerge la siguiente ventana:



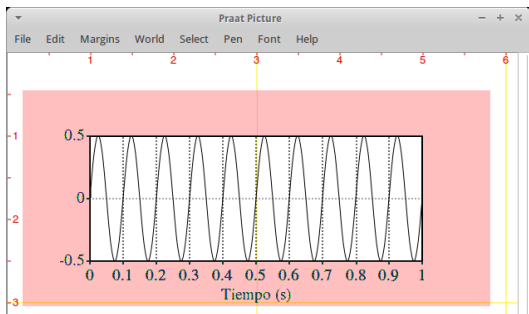
Dado que la onda tiene una duración de tan solo un segundo, no es necesario especificar límites temporales en la casilla **Time range (s)**. La amplitud está determinada con anterioridad (**Script 1a**), así que tampoco es necesario modificar **Vertical range**. La opción **Garnish** genera rótulos en inglés para la abscisa (eje temporal) y la ordenada (eje de frecuencias); asegúrese de inhabilitar esta función. Oprima **OK** y obtendrá el siguiente resultado:



3. Introduzca los rótulos. Elija en el menú **Margins** → **Draw inner box** y dibujará, como puede suponerse, un margen. Luego marque los límites temporales; para ello oprima **Margins** → **Marks** → **Marks bottom...**



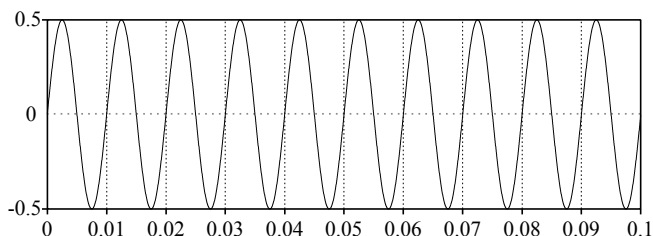
Para establecer marcas temporales en cada *ciclo* de la onda de 10Hz marque **Number of marks**: 11, y habilite las opciones **Write numbers**, **Draw ticks**, **Draw dotted lines** para dibujar los valores, diferenciarlos con una pequeña línea y trazar una línea vertical. Repita la acción en la ordenada: **Margins** → **Marks** → **Marks left...** En este caso elija **Number of marks**: 3 para establecer las marcas correspondientes a la amplitud. Finalmente, con la opción **Margins** → **Text bottom...** puede añadir texto complementario a los rótulos p. ej., **Tiempo (s)**.



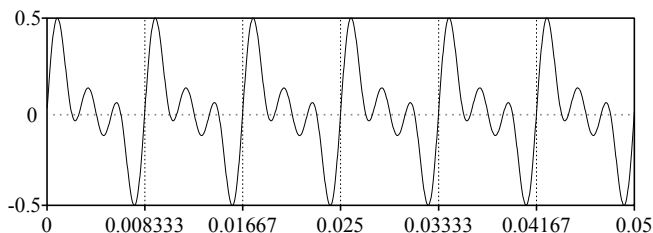


## 1.9 EJERCICIOS SUGERIDOS

1. Grabe la frase “la lima madura en mayo” utilizando diferentes tasas de muestreo p. ej., 44.100Hz, 22.050Hz, 16.000Hz y 8.000Hz. Observe detenidamente en el Editor cada frase, y anote las diferencias que puede apreciar.
2. Elabore ondas sinusoidales de 100Hz, 150Hz y 250Hz. Describa el efecto auditivo que produce el incremento de la frecuencia.
3. Calcule la frecuencia de la siguiente onda sinusoidal.



4. En fonética se acostumbra a expresar el tiempo en milisegundos (ms), sin embargo, Praat arroja la información temporal en segundos (s). Si 1.0s equivale a 1.000ms ¿cuántos milisegundos tiene el periodo de la onda anterior?
5. Con base en la información contenida en la siguiente figura, calcule el valor de  $F_0$  y la frecuencia de cada uno de los armónicos que componen la onda.



## 1.10 BIBLIOGRAFÍA RECOMENDADA

Hay una gran cantidad de obras que explican los fundamentos de la fonética acústica. Recomendando especialmente las siguientes:

Borzone, Ana María. (1980). *Manual de fonética acústica*. Argentina: Hachette. El capítulo inicial explica conceptos básicos como *sonido*, *frecuencia*, *nivel de intensidad*, *ondas complejas*. Las páginas 15-32 pueden usarse en cursos introductorios a la lingüística. Tiene un nivel básico de complejidad.

Gil, Juana. (1988). *Los sonidos del lenguaje*. Madrid: Síntesis. En el capítulo 1 (“Nociones de acústica”, páginas 11-29) se explican de manera clara conceptos como *sonido*, *onda*, *onda sinusoidal*, *onda compleja*, *análisis de Fourier*, *tonía*, *sonía*, *resonancia* y *filtros acústicos*. Este libro es recomendado para usar en cursos introductorios a la lingüística y tiene un nivel básico de complejidad.

Quilis, Antonio. (1981). *Fonética acústica de la lengua española*. Madrid: Gredos. El capítulo 2 (“Acústica del sonido”, páginas 38-62) de este libro clásico explica detalladamente los fundamentos físicos de las *ondas simples*, *ondas complejas*, *resonancia*, *filtros*. Esta obra tiene un nivel básico de complejidad y se encuentra fácilmente en las bibliotecas universitarias.

Johnson, Keith. (2003). *Acoustic and auditory phonetics*. Oxford: Blackwell. Este libro es una buena introducción a la fonética acústica y perceptiva. Una de sus características es que explica los fundamentos de la disciplina haciendo énfasis en los problemas actuales. El primer capítulo “Basic acoustics and acoustic filters” explica en detalle los conceptos de *sonido*, *ondas periódicas*, *ondas aperiódicas* y *filtros acústicos*. Este trabajo tiene un nivel intermedio de complejidad y puede usarse en cursos de licenciatura y posgrado.

Ladefoged, Peter. (1996). *Elements of acoustic phonetics*. Chicago: The University of Chicago Press. Los tres primeros capítulos de este libro son altamente recomendados como introducción a la fonética acústica y como apoyo teórico a este capítulo. La obra tiene un nivel básico de complejidad y está escrito en un estilo claro y ameno.

Lawrence, Raphael; Borden, Gloria; Harris, Katherine. (1996). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Sexta edición (2011). Baltimore: Williams & Wilkins. Esta obra está dirigida a profesionales de diferentes áreas interesados en las ciencias del habla y resulta ideal para cursos de posgrado. Las páginas 17-37 del segundo capítulo (“Acoustics”) presentan con claridad y profundidad conceptos como *movimiento armónico simple*, *la gradación de la velocidad en el movimiento armónico simple*, *la presión y movimiento de partículas*, *ondas complejas*, *ondas aperiódicas*, *frecuencia*, *tonía*, *intensidad*, *la escala de decibelios*, *sonía y velocidad del sonido*. Esta obra tiene un nivel intermedio de complejidad y está muy bien escrita.

### 1.11 ATAJO DE TECLADO

Los atajos le permitirán evitar enojosas operaciones manuales. Por ejemplo, cuando se trabaja con un archivo de audio de larga duración es necesario realizar acercamientos y alejamientos regularmente. La utilización de los botones **In** y **Out** hacen que el programa tarde demasiado en ejecutar la orden, pero los atajos **Ctrl + I**, **Ctrl + O** y **Ctrl + N** hacen que el trabajo sea ágil y podamos concentrarnos en lo que nos interesa.

**Cuadro 1.1.** Atajos de teclado para usar con Praat Objects

Opción en el menú	Teclas	Resultado
Calculator	<b>Ctrl + U</b>	Aparece un formulario para realizar una operación aritmética o resolver una fórmula. El resultado luego aparece en la ventana <b>Info</b> .
Stop playing sound	<b>esc</b>	Se detiene la reproducción del sonido actual.
Record mono Sound...	<b>Ctrl + R</b>	Abre el panel de grabación para iniciar la grabación de un sonido <b>mono</b> .
Read from file...	<b>Ctrl + O</b>	Abre archivos de audio de corta duración, archivos de texto y los objetos creados en Praat
Open long sound file...	<b>Ctrl + L</b>	Abre archivos de larga duración pero no mayores a 2Gb.
Praat Intro	<b>Ctrl + ?</b>	Se abre una ventana con el Manual de Praat.
Search Praat Manual...	<b>Ctrl + M</b>	Se abre una ventana en la que se pueden introducir términos de búsqueda de información.
Quit	<b>Ctrl + Q</b>	La orden cierra el programa. En este caso se pierden los datos que no han sido guardados.

**Cuadro 1.2.** Atajos de teclado para usar en el Editor I

Opción en el menú	Teclas	Resultado
Close	<b>Ctrl + W</b>	Cierra la ventana de edición y análisis.
Undo/Redo Cut	<b>Ctrl + Z</b>	Deshacer y rehacer.
Cut	<b>Ctrl + X</b>	Cortar una porción de sonido.
Copy selection to Sound clipboard	<b>Ctrl + V</b>	Se copia una porción de sonido en el portapapeles.
Show all	<b>Ctrl + A</b>	Muestra toda la extensión de la muestra de audio.
Zoom in	<b>Ctrl + I</b>	Acercar.
Zoom out	<b>Ctrl + O</b>	Alejar.
Zoom to selection	<b>Ctrl + N</b>	Se realiza un acercamiento a la porción de onda seleccionada.
Zoom back	<b>Ctrl + B</b>	Retroceder
Scroll page back	<b>Re Pág</b>	Desplazar hacia atrás.
Scroll page forward	<b>Av Pág</b>	Desplazar hacia adelante.
Interrupt playing	<b>esc</b>	Detener la reproducción en curso.
Copy selection to Sound clipboard	<b>Ctrl + V</b>	Copiar selección en el portapapeles.

**Cuadro 1.3.** Atajos de teclado para Praat Picture

Opción en el menú	Teclas	Resultado
Save as EPS file...	Ctrl + S	Guarda la imagen en formato .eps.
Print...	Ctrl + P	Imprime la imagen directamente.
Undo	Ctrl + Z	Sirve para deshacer acciones previas; no es posible recuperar lo borrado.
Erase all	Ctrl + E	Borra todas las imágenes de Praat Picture; no es posible recuperar lo borrado.
Praat Intro	Ctrl + ?	Se abre una ventana con el Manual de Praat.
Search Praat Manual...	Ctrl + M	Se abre una ventana en la que se pueden introducir términos de búsqueda de información en el Manual de Praat.



## 2 ANÁLISIS DEL HABLA

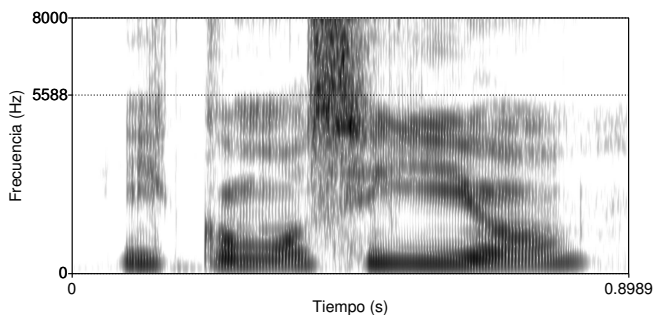
### 2.1 ESPECTROGRAMAS

Un espectrograma es la representación gráfica de las propiedades acústicas de los sonidos del habla; la frecuencia (Hz) se muestra en la ordenada, el tiempo en la abscisa y la intensidad en escala de grises. Para ver un espectrograma es necesario tener un sonido en la ventana de objetos (**Praat Objects**) y elegir **View & Edit** → **Spectrum** → **Show spectrogram**. Praat muestra por defecto un *espectrograma de banda ancha* semejante al que se muestra en la figura 2.1.

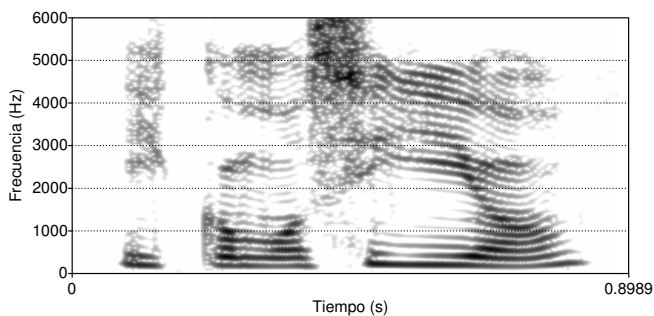
Antes de realizar cualquier análisis, deben ajustarse los parámetros de acuerdo con las características del hablante que produjo la muestra. Para ello elija en el menú **Spectrum** → **Spectrogram settings...** y defina los siguientes campos:

1. **View range (Hz)**: el programa asigna un rango entre 0.0 (Hz) y 5.000Hz, lo cual resulta adecuado para analizar hablantes masculinos. Para el habla de mujeres el rango máximo debe estar alrededor de 5.500-6.000Hz, y para analizar voz infantil puede haber información hasta 22.050Hz. El espectrograma de la figura 2.1 se extrajo del habla de una mujer: el cursor está ubicado en 5.588Hz; nótese que no hay información acústica de las vocales en las frecuencias más altas. Sin embargo, para estudiar consonantes fricativas será necesario ajustar el rango entre 0.0 y 16.000Hz. Lo fundamental es entender que las características acústicas de los sonidos que queremos estudiar y las características de los hablantes incluidos dentro del estudio determinan el rango de frecuencias adecuado para el análisis.





**Figura 2.1.** Espectrograma de banda ancha calculado con una ventana (`Window length`) de 0.004s



**Figura 2.2.** Espectrograma de banda estrecha calculado con una ventana (`Window length`) de 0.02s

2. **Window length (s):** el tamaño de la ventana de análisis se ubica entre 0.004 y 0.005s (4ms-5ms) para un *espectrograma de banda ancha*, y para un *espectrograma de banda estrecha* (figura 2.2) entre 0.02s y 0.03s (20ms-30ms). Los espectrogramas de banda ancha son ideales para estudiar los correlatos acústicos del punto y modo de articulación

de los sonidos de una lengua. En tanto, los espectrogramas de banda estrecha son ideales para estudiar características espectrales como la *frecuencia fundamental*, *formantes*, *armónicos*, *centro de frecuencia*, etc.

3. **Dynamic range (dB)**: permite ajustar el nivel de intensidad del espectrograma, lo cual se refleja en un mayor o menor contraste. El valor por defecto es de 50dB y destaca valores entre -20dB/Hz y 30dB/Hz. En tanto, los valores inferiores a -20dB/Hz aparecen en blanco (véase, [Intro 3.2 Configuring the spectrogram](#)).

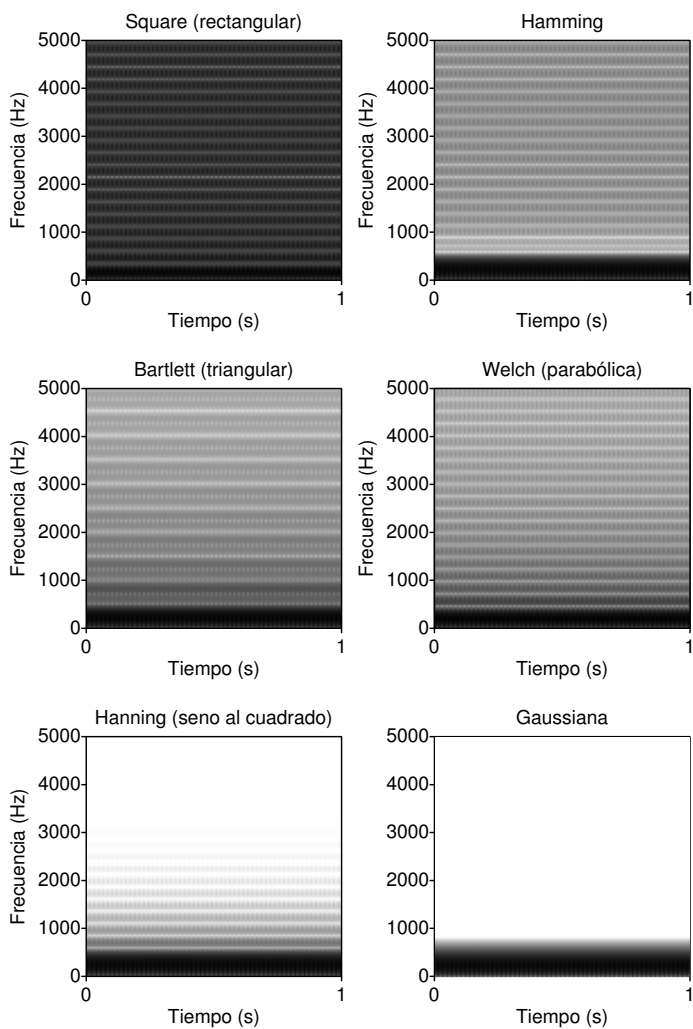
Generalmente no es necesario modificar las opciones avanzadas, pues los valores estándar están ajustados a las necesidades de la mayoría de usuarios. Sin embargo, mencionaremos cada una para dejar en claro el significado de los parámetros que estamos usando. En el menú del Editor se debe ir a  $\rightarrow$  **Spectrum**  $\rightarrow$  **Advanced spectrogram settings...** y en la ventana emergente el usuario encontrará:

1. **Number of time steps**: es el número de puntos con que se calcula el tiempo en la representación visual del espectro. Para pantallas con un ancho menor o igual a 1.200 pixeles, el valor de 1.000 es adecuado. Para pantallas grandes, los autores recomiendan optimizar la función con un valor de 1.500.
2. **Number of frequency steps**: es el número de puntos con que se calcula la frecuencia en la representación visual del espectro. Para pantallas con una altura menor o igual a 768 pixeles, el valor de 250 es adecuado. Para pantallas más altas, los autores recomiendan subir el valor hasta lograr un buen resultado.
3. **Method: Fourier** es el único método disponible.

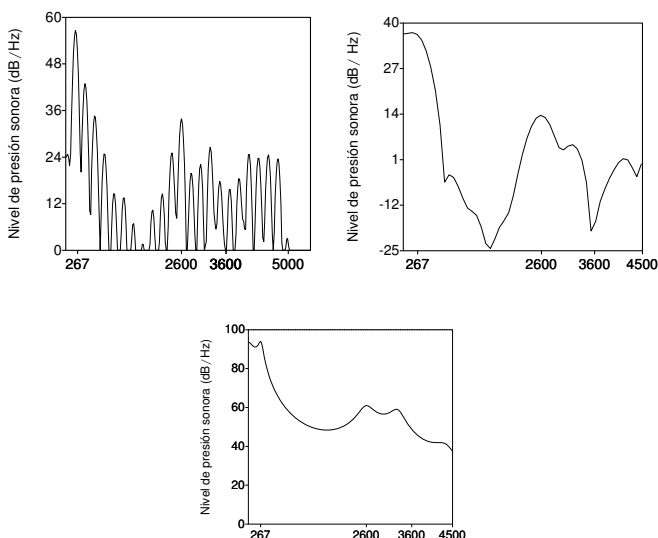
4. **Window shape**: las ventanas son funciones matemáticas que se utilizan para analizar una porción de señal. Existen diferentes formas de ventana: **square** (rectangular), **Hamming** (sinusoidal cuadrada), **Bartlett** (triangular), **Welch** (parabólica), **Hanning** (sinusoidal cuadrada) y **Gaussian** (campana). Los autores del programa recomiendan esta última, pues es la más potente y no introduce elementos ajenos a la señal como puede verse en la figura 2.3 (véase [Intro 3.2. Configuring the spectrogram](#)).
5. **Autoscaling**: la función está activada por defecto.
6. **Maximum (dB/Hz)**: los valores mayores a 100dB/Hz se representan en el espectrograma en color negro; cuando la función **Autoscaling** está activada, se toma como valor máximo el que corresponde a la parte visible del espectrograma.
7. **Pre-emphasis (dB/oct)**: el valor de +6dB por octava asegura que los formantes se vean con claridad a frecuencias altas y bajas.
8. **Dynamic compression (0-1)**: determina qué tan fuerte se ven las frecuencias débiles (y fonéticamente irrelevantes) en el espectro. El valor de 0.0 es el adecuado; si se aumenta el valor, el espectro se tornará completamente oscuro.

## 2.2 ESPECTROS

Los *espectros* son diagramas que representan las frecuencias que componen una señal compleja: en la abscisa se representa la amplitud y en la ordenada se representan las frecuencias (véase La-defoged, 1996, p.39-55). Dado que el análisis se hace sobre una porción de sonido, un espectro no presenta cambios de frecuencia en función del tiempo.



**Figura 2.3.** Tipos de ventana de análisis



**Figura 2.4.** Espectro de banda estrecha FFT (izquierda) y espectro de banda ancha FFT (derecha) y espectro continuo LPC (abajo) de una vocal [i] sintetizada

Al igual que ocurre con un espectrograma, es posible obtener *espectros de banda ancha* y *espectros de banda estrecha* (Wood, 1994-2014). Para realizarlos, Praat utiliza dos métodos: *Fast Fourier Transform* (FFT) y *Linear Predictive Coding* (LPC) (Ladefoged, 1996, p. 137-214). Para calcular un espectro FFT elija **Spectrum** → **Spectrogram settings...** y ajuste los parámetros según lo visto anteriormente. En particular, el tamaño de la ventana (Window length (s)) debe estar entre 0.02s y 0.03s (20ms-30ms) para espectros de banda estrecha, y entre 0.004 y 0.005s (4ms- 5ms) para espectros de banda ancha. Luego, se ubica el cursor en el centro de una vocal –o consonante– y se oprime **Spectrum** → **View spectral slice** (Ctrl + L).

La figura 2.4 muestra los espectros de la vocal [i] pronunciada por una mujer hablante de español bogotano. Para ver la información acústica de esta hablante, el primer espectro se calculó con una ventana de 0.02s y el segundo espectro con una ventana de 0.004s.

Para realizar un espectro continuo (LPC) es necesario hacer clic en el centro de la vocal que se desea analizar y anotar el tiempo en el que hemos ubicado el cursor. Para este fin se puede usar la tecla F6 o el comando `Query → Get cursor`. Luego vamos a `Praat Objects → Formants & LPC → to LPC (burg)`... y en la ventana emergente se deben ajustar los siguientes parámetros:

1. **Prediction order**: es el número de coeficientes de la predicción lineal. Según los creadores del programa, el valor corresponde al doble del número de picos (formantes) que se quieren analizar. Wood (1994-2014) recomienda calcular este valor con la fórmula  $n \times 2 + 2$ , donde  $n$  es el número de formantes. Por ejemplo, si tenemos un registro de audio de una voz masculina y una *Frecuencia Nyquist* de 11.000Hz, aparecerán 11 formantes, de tal suerte que el número de coeficientes será 24 ( $11 \times 2 + 2$ ). En el caso de una voz femenina, se esperan 8 formantes en 11.000 Hz y, por tanto, el número de coeficientes será de 18.

Después de los estudios realizados por Gunnar Fant en la década del sesenta, sabemos que los formantes ocurren a intervalos de  $C/2L$ , donde  $C$  es la velocidad del sonido (35.000 cm/s) y  $L$  la longitud del tracto vocal. Para un hombre con un tracto vocal de 17.5cm, se espera que ocurran formantes cada 1.000Hz (Harrington, 2012, p.84).

2. **Window length (s)**: es el tamaño de la ventana de análisis. El valor de 0.025s es adecuado para calcular espectros en muestras producidas por hombres y mujeres.
3. **Time step (s)**: es el intervalo temporal entre cada ventana de análisis. Se aconseja usar el valor por defecto.

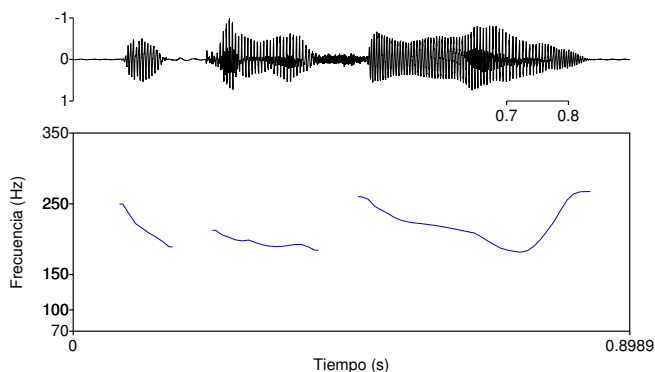
4. **Pre-emphasis frequency (Hz)**: con este parámetro se aplica un filtro que destaca las frecuencias altas. Se aconseja usar el valor por defecto.

Luego de oprimir OK se creará un objeto LPC. Para convertir el objeto en un espectro, se elige la opción **To spectrum (slice)...** en la ventana de objetos. En la casilla **Time (seconds)** se debe introducir el valor temporal del punto en que ubicamos el cursor previamente. De esta manera aseguramos que la ventana de análisis esté en un punto estable de la vocal que se desea analizar. Las demás opciones deben permanecer con el valor estándar. Oprimimos OK y se creará un objeto **Spectrum**. Para tener acceso a la representación visual del análisis se hace clic en **View & Edit**. Tenga en cuenta que el espectro FFT de banda estrecha es ideal para analizar la frecuencia fundamental y los armónicos de las vocales, mientras que el espectro de banda ancha y el espectro continuo sirven para analizar formantes.

## 2.3 FRECUENCIA FUNDAMENTAL ( $F_0$ )

Para ver la curva de  $F_0$  en el Editor se oprime **Pitch → Show pitch**. Por ejemplo, en la figura 2.5 se ven las variaciones de la curva melódica (contorno azul) del enunciado *...y ¿cómo siguió?*. Los cambios de  $F_0$  son evidentes en los sonidos sonoros (vocales, líquidas y aproximantes), mientras que en las consonantes sordas [k] y [s] se interrumpen los pulsos glotales. En el costado derecho del Editor encontrará el rango de frecuencia (p. ej., 50-350Hz). Si ubica el cursor en un punto de la curva, el programa arrojará el valor correspondiente en hercios (Hz).

Los valores de  $F_0$  dependen de la tensión y masa de los pliegues vocales, pero también del significado e intención comunicativa del enunciado, así como del hablante, su estado de ánimo y sus rasgos dialectales y sociolingüísticos. Por tal razón, es necesario ajustar los parámetros cada vez que analizamos una



**Figura 2.5.** Visualización del contorno melódico con la opción **Pitch**

voz diferente. Para realizar modificaciones elija **Pitch** → **Pitch settings...** y defina las siguientes opciones:

1. **Pitch range (Hz)**: con el rango se define el valor mínimo y máximo de  $F_0$ . El valor por defecto es 75Hz-500Hz; para voces femeninas el rango recomendado por los autores del programa es 100Hz-500Hz y para voces masculinas 75Hz-300Hz. En todo caso, estos valores son de referencia y deben modificarse de acuerdo con las características de cada hablante. De lo contrario, el programa generará una curva melódica errónea.
2. **Unit**: la unidad estándar son los **hercios (Hz)** sin embargo, para propósitos específicos es posible usar escalas de **Mels**, **Barks**, entre otras (véase Ladefoged, 1996, p.74-91).
3. **Analysis method**: el método **autocorrelation** es el adecuado para estudiar la entonación, mientras que el método **cross-correlation** es adecuado para análisis de la voz.



4. **Drawing method:**  $F_0$  se puede representar por medio curvas (**curve**), puntos (**speckles**), o por medio curvas y puntos (**automatic**).

Generalmente no es necesario modificar las opciones avanzadas. Sin embargo, mencionaremos cada una para que quede en claro el significado de los parámetros que estamos usando. En el menú se debe ir a  $\rightarrow$  **Pitch**  $\rightarrow$  **Advanced pitch settings...** y en la ventana emergente el usuario encontrará:

1. **View range (units):** en estas casillas se modifica el rango de frecuencias. Generalmente es suficiente con dejar el valor estándar i. e., 0.0 (=auto).
2. **Very accurate:** si se habilita esta opción se usa una ventana **Hanning** con la misma duración de la ventana de análisis; si se deshabilita la opción se usa una ventana **Gaussian** con el doble de duración de la ventana de análisis.
3. **Max. number of candidates:** el valor asignado es 15, y se refiere al máximo local que debe calcular el algoritmo de autocorrelación.
4. **Silence threshold:** el valor por defecto es 0.03; los tramos de señal que no superen este umbral se consideran silencio.
5. **Voicing threshold:** determina si un tramo dado es *sordo* o *sonoro*. El valor estándar es de 0.45; si se aumenta este valor hay mayor probabilidad de que los tramos analizados se consideren sordos.
6. **Octave cost:** el algoritmo usa este valor para favorecer las frecuencias altas en la detección de la periodicidad. El valor estándar es de 0.01.

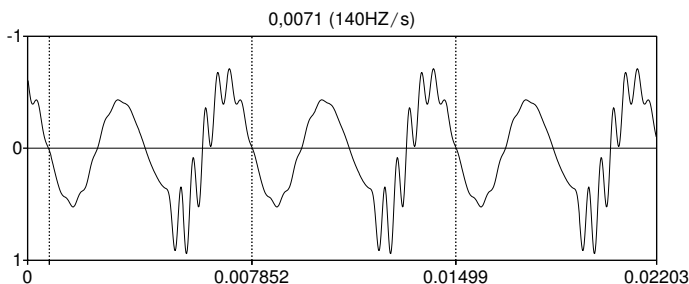
7. **Octave-jump cost**: determina la magnitud de los cambios en la curva de  $F_0$ . Al incrementar este valor se evitan caídas significativas de  $F_0$ . El valor estándar es de 0.35.
8. **Voiced/unvoiced cost**: determina la magnitud de las transiciones entre sonidos sordos y sonoros. Al incrementar este valor se reduce el número de transiciones. El valor estándar es de 0.35 (Weenink, 2014).

## 2.4 MEDICIÓN DE $F_0$

Para medir  $F_0$  existen varios procedimientos. Lo primero es seleccionar el periodo de onda (T) y calcular la frecuencia (F) con la fórmula  $F = \frac{1}{T}$  (esta información aparece en la parte superior del área seleccionada como muestra la figura 2.6). La segunda manera de hacerlo es ubicar el cursor en el punto en que queremos realizar la medición e ir a **Pitch** → **Show pitch** → **Get pitch** o simplemente pulsar F5.

La tercera posibilidad es ubicar el cursor en el lugar en que queremos realizar la medición y anotar el valor en hercios rotulado en el costado derecho de la pantalla (figura 2.7). Cuarto, podemos realizar un espectrograma de banda estrecha y ubicar  $F_0$  en la parte inferior del espectro; el valor aparecerá en el costado izquierdo de la pantalla (figura 2.7). Si se quiere mayor precisión resulta útil realizar un espectro de banda estrecha (véase la sección 2.2) y anotar el valor del primer pico espectral.

Es conveniente que el lector esté familiarizado con estos procedimientos para que esté en capacidad resolver las dudas que surgan durante el análisis, o bien para que pueda detectar errores derivados del ajuste incorrecto de **Pitch settings**.



**Figura 2.6.** Cálculo de  $F_0$  con la fórmula  $F = \frac{1}{T}$  realizado sobre una vocal sintética [i] con  $F_0$  de 140Hz. El resultado aproximado del procedimiento es  $140Hz = \frac{1}{0,007156}$

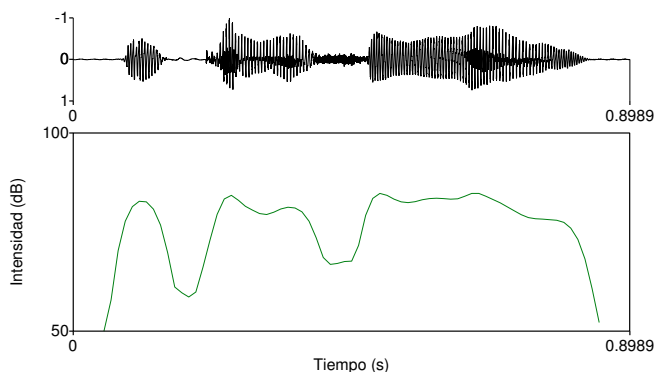


**Figura 2.7.** Cálculo de  $F_0$  con un espectrograma (izquierda) y con la opción automática (derecha). La primera opción es menos precisa pues depende de la ubicación exacta del cursor

## 2.5 INTENSIDAD

La curva de intensidad muestra las variaciones de presión de un sonido expresadas en decibelios (dB). Sin embargo, los estímulos auditivos provocados por los cambios de intensidad no se relacionan con cambios de presión de igual magnitud. Por esta razón se dice que la percepción de la intensidad es *no lineal* (Johnson,

2006, p.49). Para ver la intensidad de una señal de habla se debe habilitar la opción **Show intensity** en el Editor, y el programa arrojará una representación como la que se observa en la figura 2.8. Para ajustar los parámetros de análisis vaya a **Intensity** → **Intensity settings...** De acuerdo con los autores ([Configuring the intensity contour](#)), los principales parámetros son:



**Figura 2.8.** Curva de intensidad

1. **View range:** el rango estándar (50-100dB) se ajusta bien al umbral de percepción de los sonidos del habla, pero el usuario puede ajustarlo en caso de ser necesario.
2. **Averaging method:** “se refiere al método que Praat usa para promediar la intensidad. Si el método es **energy**, el valor en decibelios que arroja el programa se basa en la media de la potencia (expresada en  $\text{Pa}^2/\text{s}$ ) entre los límites del rango ( $t1$  y  $t2$ ). Si el método es **dB**, el valor arrojado corresponde a la media de la curva de intensidad. Si el método es **sones**, el valor arrojado se expresa en unidades

cuya distancia está relacionada con ciertas propiedades de procesamiento del oído humano” ([Intensity Get mean...](#) en el manual oficial de Praat).

3. **Subtract mean pressure:** al habilitar esta opción se extrae la media de la presión sonora con una ventana gaussiana.

## 2.6 MEDICIÓN DE LA INTENSIDAD

La manera más sencilla de medir la intensidad de una señal es seleccionar un área de análisis o ubicar el cursor en un punto en el tiempo y leer el valor que se lista en el costado derecho de la pantalla. Una manera más exacta de realizar la medición es utilizar la opción **Intensity listing** o presionar la tecla F8, en cuyo caso obtendremos un archivo de texto (**Praat info**) con los cambios de intensidad que se suceden en la porción de señal.

Con **Get intensity** se obtiene el valor medio de la intensidad del área seleccionada. Si quiere conocer el valor de la intensidad mínima puede elegir **Get minimum intensity** u oprimir **Ctrl + F8**; si quiere conocer la intensidad máxima de una selección elija **Get maximum intensity** u oprima **Mayús + F8**.

## 2.7 FORMANTES ( $F_n$ )

Los *formantes* son picos espectrales que permiten la discriminación e identificación de las vocales y algunas consonantes de las lenguas del mundo. En especial, los dos primeros formantes (abreviados F1 y F2) arrojan información sobre la altura de la lengua, su posición en la dimensión anterior/posterior del tracto vocal y el grado de redondeamiento de los labios. El tercer formante (F3) tiene importancia para la percepción de las vocales anteriores (Liljencrants y Lindblom, 1972).

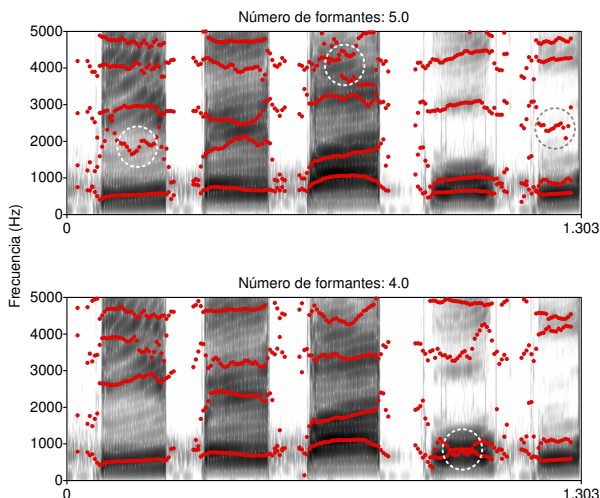
En términos generales, el valor de F1 aumenta cuando se abre la cavidad oral –como en la vocal [a]– y el valor disminuye cuando la lengua se eleva y forma una constricción como en [i] o [u]. En tanto, F2 es alto cuando la lengua forma una constricción en la región palatal, y el valor disminuye a medida que la lengua se desplaza hacia la parte posterior de la boca o cuando hay redondeamiento de los labios. Tal es el caso de las vocales [o] y [u] del español (para mayor detalle véase Fant, 1960; Ladefoged, 1996; Johnson, 2006).

Para ver los formantes en Praat habilite la opción **Show spectrogram**. Luego vaya a **Formant** → **Show formants**. En ese punto aparece un contorno rojo destacando los formantes de la muestra que se está analizando (véase la figura 2.9).

Sin embargo, para obtener mediciones correctas de los formantes es necesario estar familiarizado con los parámetros de análisis. Para acceder a las opciones de configuración es necesario ir a **Formant** → **Formant settings...** En la ventana emergente encontrará:

1. **Maximum formant (Hz)**: el valor estándar es 5.500Hz y sirve para analizar cinco formantes a intervalos de 1.100Hz en habla femenina. Para el habla de hombres el valor recomendado es 5.000Hz y supone un intervalo entre formantes de 1.000Hz. Estos valores estándar deben entenderse como una referencia para el usuario, de manera que es necesario ajustar el valor de acuerdo con las características de cada hablante.
2. **Number of formants**: el valor asignado es 5 formantes. Se sugiere aumentar a 6 si se ven menos formantes de los esperados, y reducirlo a 4 o 3 si se ven más formantes (Styler, 2012). Por ejemplo, en la figura 2.9 se muestran las sílabas ['pi, 'pe, 'pa, 'po, 'pu] pronunciadas por una mujer del Pacífico de Colombia. En este caso se ajustó el análisis en 5 formantes, pero esto introdujo falsos formantes en

la primera, tercera y quinta sílabas. Este problema lo solucioné reduciendo el valor a 4.0, lo cual resultó adecuado para analizar cuatro vocales. En el caso de la vocal posterior de [pu] fue necesario hacer una medición aparte considerando 5 formantes. Nota: siempre que se modifique **Number of formants** debe ajustarse el valor de **Maximum formant**.



**Figura 2.9.** Error de cálculo de los formantes

3. **Window length (s)**: si se asignan 0.025s se aplica una ventana gaussiana de 0.050s, lo cual resulta efectivo en la mayoría de los casos.
4. **Dynamic range (dB)**: 30dB/Hz es un valor adecuado para la visualización de formantes.
5. **Dot size(mm)**: si se aumenta el valor, los formantes se ven representados con puntos más intensos y definidos. Para

elaborar la figura 2.9, por ejemplo, se utilizó un grosor de 1.5 en el contorno de los formantes.

En **Formant** → **Advanced formant settings...** se especifica el método que Praat utiliza para calcular los formantes (**Burg**). Si el valor del **Pre-emphasis** es 50.0, se ignoran las frecuencias inferiores y las demás incrementan +6dB por octava. Así, las frecuencias a 100Hz se incrementan +6dB, a 200Hz se incrementan +12dB, etc., (**Sound: To Formant (burg)**).

## 2.8 MEDICIÓN DE $F_n$

Si tenemos en cuenta los parámetros expuestos, podremos ver el contorno de formantes superpuesto sobre el espectrograma y realizar mediciones en diferentes puntos del tiempo con relativa rapidez. La manera más sencilla de medir los formantes es ubicar el cursor en un punto estable de una vocal (i. e., sin movimientos significativos del contorno del formante) y leer el valor rotulado en color rojo al costado izquierdo de la pantalla. Por ejemplo, en la señal que se muestra en la figura 2.9, la vocal [e] tiene un F2 de 2.296Hz a los 0.422106 segundos del inicio de la grabación.

Una manera más precisa de realizar la medición es ir al menú **Editor** y elegir **Formant** → **Formant listing**. Así, Praat arrojará una tabla con el valor de los cuatro primeros formantes en el punto del tiempo en que está ubicado el cursor:

Time.s	F1.Hz	F2.Hz	F3.Hz	F4.Hz
0.422106	751.086713	2336.408885	2886.893659	4578.197808

Para conocer el valor de un formante de manera rápida, tan solo es necesario oprimir las teclas F1 o F2. En nuestro ejemplo, el valor de F2 obtenido con este procedimiento es el siguiente:

2336.408885146047 Hz (nearest F2 to CURSOR)



El dato concuerda con el arrojado por **Formant listing**, pero difiere en 40Hz de la medición manual ( $F_2 = 2296\text{Hz}$ ). Esto significa que debemos evitar el método manual para analizar estadísticamente datos destinados a la divulgación científica. La medición manual puede resultar útil para quienes conocen los valores promedio de las vocales de una lengua y quieren despejar pequeñas dudas durante una transcripción o una breve inspección de la señal; para el aprendiz la medición manual puede resultar desastrosa y, por tanto, es un hábito de análisis que recomiendo evitar.

La manera tradicional de medir formantes es calcular el espectro con las técnicas *Fast Fourier Transform* (FFT) y *Linear Predictive Coding* (LPC). Para implementarlas en Praat basta con seguir las indicaciones dadas en la sección 2.2. En caso de que las instrucciones no hayan funcionado y el espectro resultante no muestre claramente los picos, debe relacionar la tasa de muestreo (**sampling frequency**) del archivo que está usando con el rango de frecuencias en el que se encuentra la información que quiere medir. Para el caso de las vocales, la frecuencia máxima relevante está alrededor de 5.500Hz y, por tanto, la frecuencia de muestreo óptima es 11.000Hz. Si desea calcular el espectro de consonantes fricativas o la explosión de las oclusivas sordas (*Stop burst*) es recomendable usar la técnica FFT y remuestrear la señal a 16.000Hz y 20.000Hz respectivamente (Ladefoged, 1996, p.178). La figura 2.10 muestra espectros LPC para las cinco vocales del español (datos de Martínez-Celdrán, 1991, p.18-19)

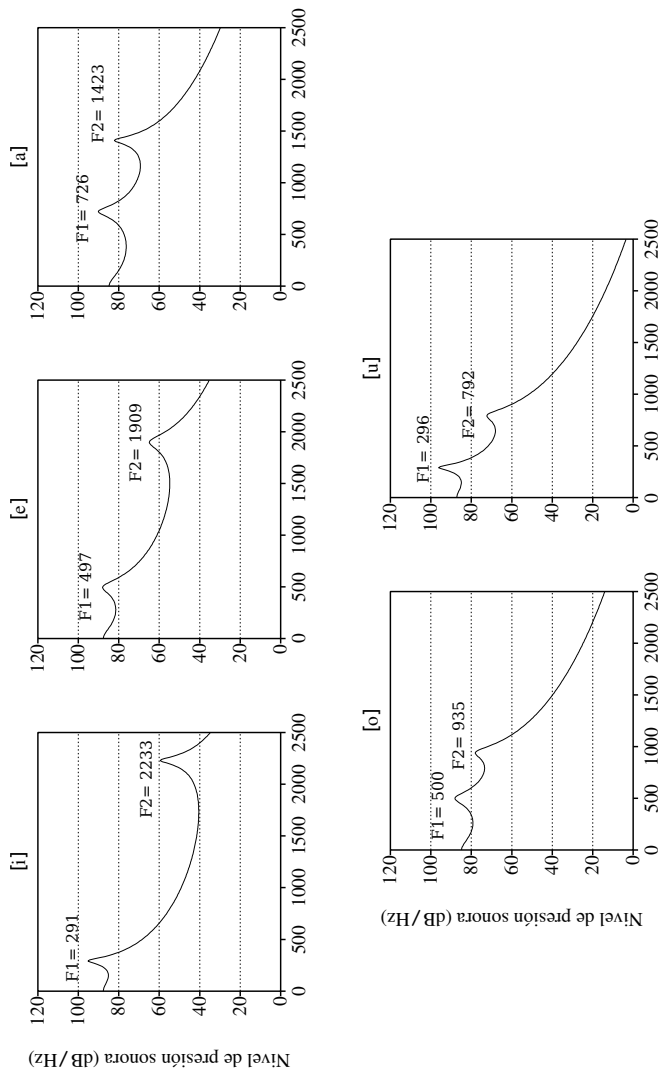
Si no conoce la frecuencia de muestreo del archivo de audio debe seleccionarlo en la ventana de objetos y oprimir el botón **Info** en la parte inferior izquierda. Luego, una ventana arrojará la información de la señal, tal como se explicó en el primer capítulo. Si es necesario cambiar la frecuencia de muestreo, en el menú dinámico debe elegir **Synthesize** → **Convert** → **Resample...**, introducir la nueva tasa y oprimir **OK**. De es-

ta manera se creará un objeto de audio que podrá utilizar para calcular el tipo de espectro que desee. El objetivo del procedimiento es mejorar el cómputo que hace Praat de las frecuencias: al reducir la tasa de muestreo se reduce el intervalo entre los componentes (Ladefoged, 1996, p. 177–178). En todo caso, Boersma y Weenink (2014) no recomiendan usar *Fast Fourier Transform* y *Linear Predictive Coding* para medir formantes en Praat. En cambio, aconsejan usar la opción **To Formant (burg)**... que está disponible en la ventana de objetos. En la siguiente sección se explica cómo usarla.

## 2.9 TO FORMANT (BURG...)

Para medir los formantes desde la ventana de objetos debe seleccionar el archivo de audio, oprimir el botón **Analyse spectrum** y la opción **To Formant (burg...)**. Como resultado, saldrá una ventana con las opciones explicadas en la sección 2.7. Una vez se han llenado los campos se creará un objeto **Formant**, el cual contiene información sobre el contorno de formantes del archivo. Para obtener el valor de un formante solo debe oprimir **Query → Get value at time...** e introducir el punto en el tiempo en que desea hacer la medición (p. ej., 0.55s) y el número del formante (p. ej., F1, F2).

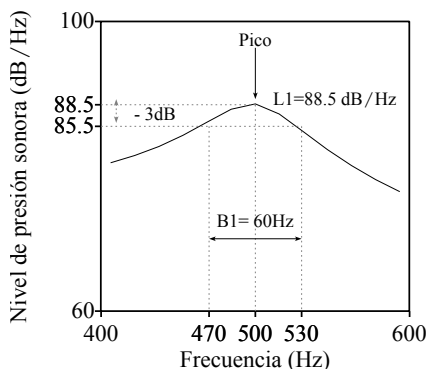
Con la opción **To Formant (burg...)** no es necesario remuestrear la señal de audio. Por el contrario, si calcula los formantes por medio de un espectro con los métodos *Fast Fourier Transform* o *Linear Predictive Coding* debe remuestrear la señal. Como se vio arriba, si el locutor es una mujer, los formantes estarán cada 1.100 Hz, así que tendrá que remuestrear la señal a 11.000 Hz para contar con información óptima entre 0 y 5.500 Hz. Sin embargo, **To Formant (burg...)** es una alternativa más práctica para incluir en scripts para extraer formantes.



**Figura 2.10.** Espectros LPC de los dos primeros formantes de las vocales del español (datos de Martínez-Celdrán, 1991, p.18-19)

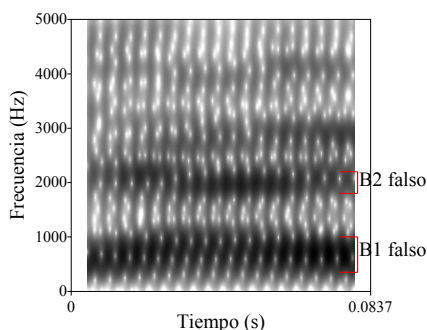
2.10 ANCHO DE BANDA ( $B_n$ )

Como se ha visto, los formantes se representan en un espectrograma como contornos que varían a lo largo del tiempo. Dichos contornos también varían dentro de un rango de frecuencias o, más específicamente, tienen un *ancho de banda* ( $B_n$ ). Acústicamente, el ancho de banda se define como el rango de frecuencias efectivas de un resonador (Ladefoged, 1996, p.68) o la diferencia de frecuencias entre dos puntos adyacentes a un pico espectral (Fant, 1956, p.111). De acuerdo con Stevens (1998), el ancho de banda está relacionado con las pérdidas acústicas del tracto vocal producidas por la viscosidad y la conducción del calor.



**Figura 2.11.** Ancho de banda ( $B_1$ ) y nivel del formante ( $L_1$ )

En la figura 2.11 se muestra el espectro del primer formante de una vocal central media [ə]. El valor del formante es 500Hz y su nivel de intensidad ( $L_n$ ) es de 88.5dB/Hz. Para determinar  $B_1$  se ubicaron dos puntos del espectro con un nivel de intensidad de -3dB por debajo del nivel alcanzado en el pico espectral, los cuales corresponden a 470Hz y 530Hz. Como puede verse, la diferencia entre estos puntos (i. e., ancho de banda) es de 60Hz.



**Figura 2.12.** Bandas del espectrograma ( $B_n$  falso) de un espectrograma de banda ancha

## 2.11 MEDICIÓN DE $B_n$

Para medir el ancho de banda en Praat debe tener un objeto **Formant** en la ventana de objetos (véase el apartado anterior), y luego oprimir **Query- → Get bandwidth at time...** Si tiene una señal en el editor debe ir a **Formant → Get first bandwidth, Get second bandwidth, Get third bandwidth o Get fourth bandwidth**. Como advierte Fant (1956, p.111), “el ancho de banda no debe confundirse con el ancho de las bandas en el espectrograma, pues estas dependen esencialmente de los parámetros de análisis”<sup>1</sup>. La figura 2.12 muestra las bandas del espectrograma ( $B_n$  falso) de un espectrograma de banda ancha calculado con las opciones estándar de Praat.

## 2.12 CREACIÓN DE VOCALES CON VOWELEDITOR

Una manera para entender la importancia de F1 y F2 en la determinación de la cualidad vocálica es creando nuestros pro-

---

<sup>1</sup>La traducción es mía.

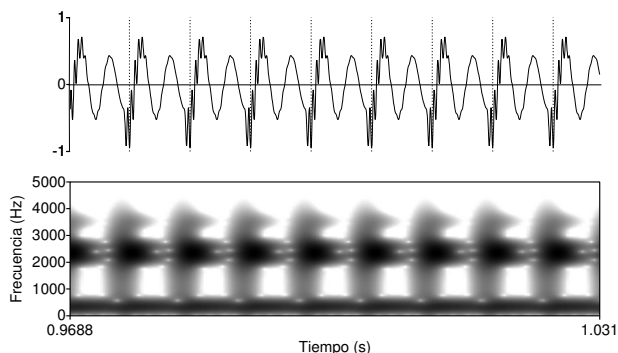
prios estímulos. Para ello, Praat ofrece una herramienta llamada **VowelEditor**, a la cual podemos acceder desde la ventana de objetos con los comandos **New** → **Sound** → **Create Sound from VowelEditor...**

**VowelEditor** muestra el espacio acústico formado de la correlación de los dos primeros formantes ( $F_1 \times F_2$ ), y arroja una reproducción de las cualidades correspondientes a cada punto del plano. Para usar la herramienta debe hacer clic sobre una vocal y el programa reproducirá el sonido correspondiente. También puede crear un sonido que muestre la transición entre dos o más cualidades. Con el botón **Play** se reproduce la vocal, **Reverse** cambia el sentido de una trayectoria y con **Publish** se obtiene un nuevo objeto en la ventana principal.

En la parte inferior del panel hay varias opciones que permiten mejorar la calidad de los estímulos. Por ejemplo, se puede modificar la duración total (**Duration**), la duración de una trayectoria (**Extend**),  $F_0$  (**Start F0**) y la caída de  $F_0$  (**F0 slope**). Además, en el menú principal de **VowelEditor** se pueden añadir valores para  $F_3$  y  $F_4$ , con lo cual se logrará mayor naturalidad en el resultado. Si le resulta necesario introducir estos formantes, los comandos son **Edit** → **Set F3 & F4...** Invito al lector a que explore auditivamente el editor de vocales y otras opciones que ofrece su menú.

## 2.13 PULSOS GLOTALES

La última opción en el menú del Editor es el análisis de los pulsos glotales. Estos se muestran delimitados por líneas verticales con los comandos **Pulses** → **Show pulses**. Para que el programa realice correctamente los cálculos se debe ir a **Pitch** → **Pitch settings** y elegir el método **cross-correlation**, pues **autocorrelation** se aplica al análisis de la curva melódica. Los autores del programa recomiendan ajustar el rango de  $F_0$  entre 50-200Hz para analizar voz patológica (véase, **Voice**).



**Figura 2.13.** Pulsos glotales o Pulses

Si usted desea analizar los pulsos glotales debe seleccionar el área de interés y realizar un acercamiento con el botón **sel** (parte inferior izquierda del Editor) o con el atajo **Ctrl + N** (figura 2.13). Luego debe elegir **Pulses** → **Voice report** y luego obtendrá un listado con la siguiente información:

1. **Time range of SELECTION:** indica el tiempo de inicio y final de la selección.
2. **Pitch:** obtenemos medidas de centralización (media, mediana, moda) y la desviación estándar de  $F_0$ .
3. **Pulses:** el número de pulsos en la selección, el número de periodos, la media de los periodos y su desviación estándar.
4. **Voicing:** indica el número de tramos de análisis en los que se mantiene la fonación, el número de interrupciones (*voice breaks*) y su magnitud.
5. **Jitter:** computa los periodos de los ciclos para determinar la perturbación de  $F_0$ .

6. **Shimmer**: es una medida que sirve para determinar la perturbación de la amplitud de cada ciclo.
7. **Harmonicity**: determina el grado de la periodicidad acústica. Praat arroja datos sobre la media de la relación armónico-ruido (*harmonics-to-noise ratio*) y la media de la relación ruido-armónico (*noise-to-harmonics ratio*).

## 2.14 DURACIÓN

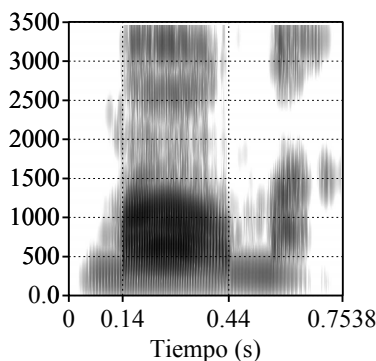
La duración es un parámetro que se tiene en cuenta en los análisis fonéticos por su importancia en la determinación de fenómenos como el *acento*, la *cantidad* y el *timbre* vocálico. Por ejemplo, en la lengua uitoto o murui la distinción entre vocales breves y largas permite distinguir una parte importante del vocabulario p. ej., *mooma* 'padre' (figura 2.14) y *diie* 'ceniza' (Becerra y Petersen, 2003). En lenguas como el inglés y el sueco también se establecen diferencias de cantidad pero, además, las vocales átonas suelen ser breves y realizarse como una vocal central media [ə] o una vocal central cerrada [ɪ] (Lindblom, 1963).

La duración está correlacionada con el *tempo* o *velocidad de elocución*, el tipo de sonido (i. e., la duración intrínseca), el *contexto fonético*, la frecuencia de ocurrencia y el tipo palabra (Pluy-maekers et al., 2005). A nivel entonativo, las vocales son largas al final de las declarativas, y en las frases exclamativas los sonidos generalmente involucran mayor tiempo de producción. El lector puede remitirse a Lehiste (1977) para una presentación detallada de la *duración* y su correlación con los diferentes sonidos y niveles de análisis fonético.

## 2.15 MEDICIÓN DE LA DURACIÓN

La duración puede medirse desde la ventana de objetos, desde el menú del Editor o directamente desde la señal. La manera más





**Figura 2.14.** Vocales largas en la palabra /mo:ma/ ‘padre’ del uitoto (*murui-bue*)

sencilla, por supuesto, es seleccionar una porción de sonido en el editor y anotar los valores que arroja Praat. También puede ir al menú del Editor y elegir la opción **Query**. Así podrá averiguar el tiempo en que inicia un sonido (**Get start of selection**), el tiempo de finalización (**Get end of selection**), la duración total (**Get selection Length**) y el punto en el tiempo en que está ubicado el cursor (**Get cursor**).

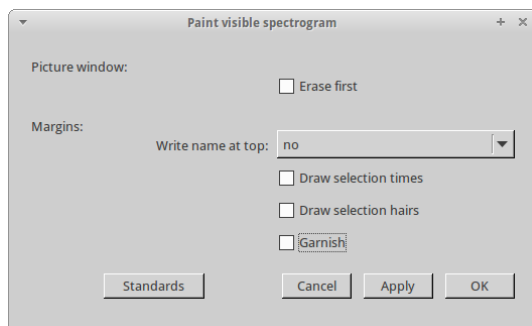
Para extraer la duración de una porción de sonido desde la ventana de objetos tan solo debe seleccionarlo, oprimir el botón **Query- → Query time domain**, y podrá extraer el tiempo en que inicia un sonido (**Get start time**), el tiempo de finalización (**Get end time**) y la duración total (**Get total duration**). Esta alternativa de análisis es particularmente útil cuando se tiene un sonido segmentado en un objeto **TextGrid** (véase la sección 3.2) y es necesario extraer valores temporales de sonidos individuales.

## 2.16 DIBUJANDO UN ESPECTROGRAMA

Aprender a crear espectrogramas y rotularlos es una tarea importante para presentar de manera adecuada nuestros datos en artículos o trabajos de clase. La cantidad de información que debe tener un espectrograma depende siempre de las características acústicas que se quieran resaltar. Por ejemplo, si el interés es mostrar el valor de los formantes de una vocal en un rango entre 0-5.000Hz, entonces será suficiente con hacer marcas cada 1.000Hz y ajustar el rango de frecuencias del sonido. Si el interés es mostrar la duración de un ‘segmento’ dentro de una preferencia, entonces el eje de la abscisa deberá tener marcas que permitan al lector hacer el contraste.

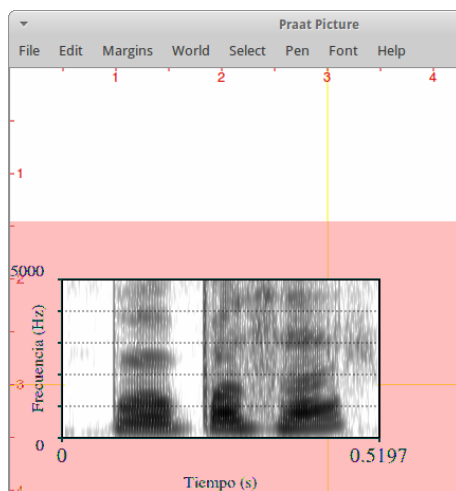
Los fonetistas llevan décadas usando espectrogramas, así que la manera como se presentan estos datos varía enormemente entre los diferentes autores. De acuerdo con mi experiencia como lector, sugiero dibujar un espectrograma en Praat siguiendo los siguientes pasos.

1. Después de elegir cuidadosamente el sonido que utilizará como ejemplo, debe hacer una selección amplia que permita incluir un oscilograma o forma de onda en la parte superior. Luego oprima **Spectrum** → **Paint visible spectrogram...** y obtendrá la siguiente ventana:



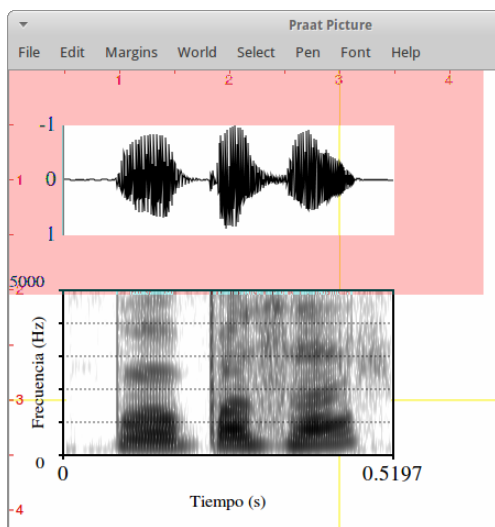
En la parte superior la opción **Erase first** le permite cerciorarse de borrar imágenes previas. En el margen superior puede dibujar el nombre del objeto (**write name at top**), dibujar marcas temporales (**Draw selection times**) con líneas verticales (**Draw selection hairs**) dentro de una cuadrícula y etiquetas en inglés (**Garnish**). Sugiero deshabilitar estas opciones. Oprima OK y tendrá un espectrograma sin rótulos.

2. Para dibujar un recuadro oprima **Margins** → **Draw inner box**. Después incluya el tiempo de inicio y finalización del sonido con **Margins** → **Marks bottom...** En la ventana emergente coloque dos marcas (**Marks bottom...**) y deje las demás opciones como se encuentran. Para incluir los valores de frecuencia en la ordenada elija **Margins** → **Marks left...** En la imagen que se muestra abajo, por ejemplo, se incluyeron seis marcas para dividir la escala en cinco partes distribuidas cada 1.000Hz.



Para que sus rótulos se distingan puede incluir marcas (**Draw ticks**), los valores numéricos (**Write numbers**) y trazar líneas a lo largo de la imagen (**Draw dotted lines**). Añada texto con las opciones **Margins** → **Text left...** o **Margins** → **Text bottom...**

3. Finalmente, es conveniente complementar el espectrograma con la forma de onda. Para dibujarla, debe elegir desde el editor **File** → **Draw visible sound...** En la ventana emergente deshabilite los rótulos (**Garnish**) y todas las opciones de dibujo (**Draw selection times**, **Draw selection hairs**). La línea vertical que aparece en la imagen la dibujé con **World** → **Draw line...** y ajusté los parámetros así: **From x: 0.0 To x: 0.0**, **From y: 1.0 To y: -1.0**. Finalmente, rotulé la amplitud con las opciones **Margins** → **Marks left...** e introduje tres marcas.



## 2.17 BIBLIOGRAFÍA RECOMENDADA

Borzone, Ana María. (1980). *Manual de fonética acústica*. Buenos Aires: Hachette. En el cuarto capítulo la autora explica conceptos básicos como *formante, nivel de intensidad, ancho de banda, frecuencia fundamental, intensidad, duración*, las relaciones entre las frecuencias del formante y la forma de la cavidad oral, y cierra con una presentación de las características acústicas de las vocales del castellano. En el quinto capítulo el lector encontrará una presentación de las características acústicas de sonidos consonánticos (*oclusivas orales, oclusivas nasales, africadas, fricativas, laterales*).

Johnson, Keith. (2003). *Acoustic and Auditory Phonetics*. Oxford: Blackwell. El quinto capítulo es una introducción amable a la *Teoría acústica de la producción del habla*. El sexto capítulo discute las características acústicas de las vocales y los modelos de tubos. El séptimo capítulo trata sobre las fricativas y la teoría cuántica del habla. Los capítulos finales tratan sobre las consonantes *oclusivas, africadas, nasales y laterales*.

Ladefoged, Peter. (1996). *Elements of Acoustic Phonetics*. Chicago: The University of Chicago Press. Los capítulos *Análisis de ondas* (4), *Resonancia* (5) y *La producción del habla* (7) abordan de manera clara y breve los aspectos teóricos relacionados con este capítulo del manual.

Lawrence, Raphael; Borden, Gloria; Harris, Katherine. (1996). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Sexta edición (2011). Baltimore: Williams & Wilkins. El quinto capítulo "The Articulation and Acoustics of Vowels" presenta la *Teoría acústica de la producción del habla* y la relación *articulación/señal acústica* a propósito de las vocales del inglés. El sexto capítulo está dedicado a las consonantes (*nasales, fricativas, oclusivas*).

Martínez Celdrán, Eugenio y Fernández Planas, Ana María. (2007). *Manual de fonética española*. Barcelona: Ariel. En este trabajo el lector puede encontrar información sobre la articulación y las propiedades acústicas de los sonidos del español. El capítulo introductorio trata sobre la interpretación de electropalatogramas e imágenes con información acústica. El resto del libro está dedicado a estudiar en capítulo aparte las *obstruyentes* (oclusivas, aproximantes espirantes y africadas), *sonantes* (nasales, laterales y róticas) y las *vocales*. El libro abarca aspectos teóricos centrales y los datos teóricos y experimentales son actuales. Puede usarse en cursos de *fonética y fonología española* y como apoyo teórico de los temas discutidos aquí.

Quilis, Antonio. (1981). *Fonética acústica de la lengua española*. Madrid: Gredos. En esta obra el estudio de las características acústicas de los sonidos del castellano se aborda en cinco capítulos dedicados a las *vocales*, *oclusivas*, *fricativas*, *africadas* y *líquidas*, respectivamente. Sugiero acompañar esta lectura con bibliografía actual y contrastar el uso de la terminología con autores actuales.

Stevens, Kenneth. (1998). *Acoustic Phonetics*. Cambridge, MA: The MIT Press. Esta obra de nivel avanzado es la mayor autoridad en el tema. El libro está dividido en diez capítulos en los que se exponen los principios de la *Teoría acústica de la producción del habla*. El lector encontrará información sobre anatomía y fisiología del habla, los mecanismos de excitación laríngea, las propiedades de resonancia del tracto vocal, el procesamiento auditivo de los sonidos, y capítulos individuales dedicados a las consonantes oclusivas (*transiciones de formantes*), consonantes obstruyentes, sonantes, vocales y un capítulo sobre la influencia del *contexto consonántico* sobre las propiedades espectrales de las vocales.

## 2.18 RECURSOS RECOMENDADOS

En Internet se encuentran disponibles varios manuales de Praat y conjuntos de **Scripts** creados por diferentes investigadores. Recomendando especialmente los siguientes:

Hirst Daniel. *Scripting Praat: a tutorial for beginners*. Este manual explica los comandos y las estructuras básicas para escribir un script. El manual está dividido en dos partes cortas y fáciles de comprender y enseña cómo medir la *duración* de una señal etiquetada en Praat. ENLACE:

[http://uk.groups.yahoo.com/group/praat-users/files/ Daniel\\_Hirst/Scripting\\_Praat/](http://uk.groups.yahoo.com/group/praat-users/files/Daniel_Hirst/Scripting_Praat/)

Lenne, Mieta. *The Speech Corpus Toolkit for Praat*. En esta página la autora pone a disposición un conjunto de scripts para realizar tareas como analizar *formantes*,  $F_0$ , la *duración*, *dibujar cuadros de formantes y espectros*, etc. Para comprender el uso de los recursos es necesario que el lector tenga dominio de las funciones básicas de Praat. ENLACE: <http://www.helsinki.fi/~lennes/praat-scripts/>

Llisterri, Joaquim. *El análisis acústico del habla mediante Praat*. Este manual aborda los temas presentados en este capítulo i. e., análisis oscilográfico, espectrográfico, espectral (FFT y LPC), análisis del espectro medio (LTAS). Además, trata sobre segmentación y etiquetado, manipulación del habla, creación de estímulos sintetizados y uso de scripts. Recomendando especialmente el apartado *La extracción automática de los valores de los formantes*. ENLACE: [http://liceu.uab.es/~joaquin/phonetics/fon\\_Praat/Praat.html](http://liceu.uab.es/~joaquin/phonetics/fon_Praat/Praat.html)

Román, Domingo. *Manual de Praat*. En capítulos cortos y directos el autor aborda los siguientes temas: grabación, edición de señales, etiquetas, espectrogramas, análisis temporal, formantes, análisis del tono, manipulación del tono y crea-

ción de gráficos. El capítulo sexto, dedicado a los *formantes*, es un buen complemento de los temas presentados aquí, pues incluye un script para tabular los formantes de una muestra previamente etiquetada. ENLACE: [http://domingo-roman.net/manual\\_analisis\\_acustico.html](http://domingo-roman.net/manual_analisis_acustico.html)

Styler, Will. *Using Praat for Linguistic Research*. Este tutorial explica detalladamente el uso del programa y da información sobre la configuración de Praat para realizar análisis acústicos. El texto está bien organizado y tiene una buena variedad de contenidos. Este manual aborda los contenidos presentados aquí. Las personas interesadas en mediciones de la *nasalidad* y el *tipo de fonación* encontrarán mediciones útiles. ENLACE: <http://savethevowels.org/praat/>

UCLA Phonetics Laboratory. *Praat script resources*. Esta página reúne decenas de **scripts** que incluyen: gestión de archivos de audio y objetos **TextGrid**, segmentación, manipulación del habla, creación de imágenes, etc. ENLACE: <http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/praat.html>

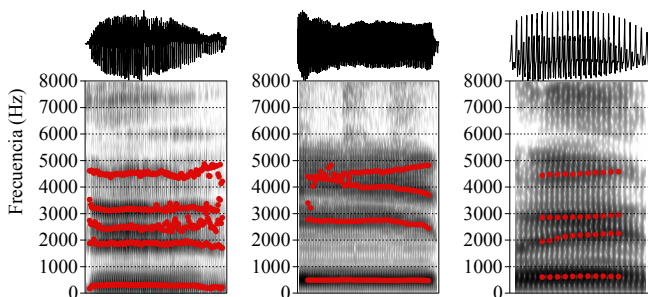
van Lieshout, Pascal. *Praat tutorial*. Este tutorial explica las funciones básicas de Praat; es recomendable para usuarios de Windows sin conocimientos avanzados en informática. ENLACE: <http://oraldynamicslab.ca/>

Wood, Sydney. *Praat for beginners*. Al igual que los trabajos mencionados, este explica cómo usar el programa para realizar análisis acústicos del habla. Los consejos del autor para realizar figuras y rotularlas producen resultados sencillos y bellos. ENLACE: <http://swphonetics.com/>



## 2.19 EJERCICIOS SUGERIDOS

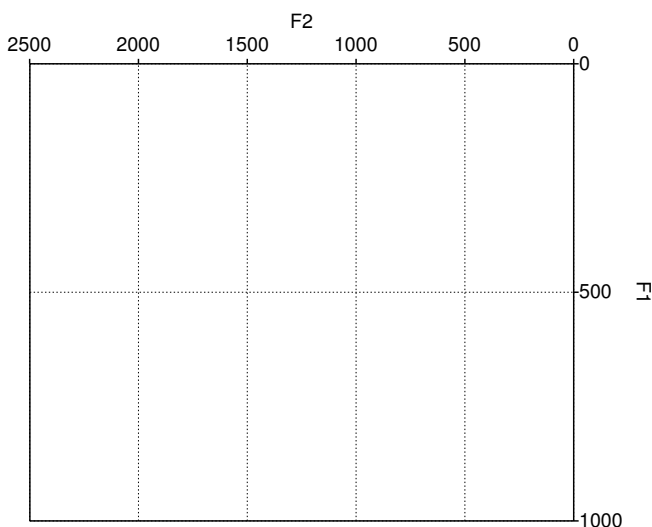
1. La siguiente figura muestra tres sonidos pronunciados por una mujer, un hombre y un niño. Teniendo en cuenta las tendencias del *rango de frecuencias de la voz humana*, identifique cuál espectrograma corresponde a cada individuo. Justifique su respuesta.



2. Grabe su producción de las vocales del español /i e a o u/. Para este fin, grabe frases marco del tipo ‘diga  $x$  dos veces’, donde  $x$  es una palabra con acento en la primera sílaba p. ej., *casa*, *queso*, *kilo*, *cura*. Realice un espectrograma de banda estrecha de cada vocal, y mida los cinco primeros armónicos desde el Editor. Luego realice un espectro de banda estrecha y mida nuevamente los armónicos. ¿Qué diferencias puede apreciar entre los dos métodos?

3. Grabe un enunciado constituido por sonidos sonoros p. ej., *la luna madura en mayo* y calcule el valor *medio*, *mínimo* y *máximo* de su  $F_0$ .

4. Grabe las palabras *saco*, *seco*, *cita*, *soga* y *susto* en una frase marco del tipo ‘diga  $x$  dos veces’, donde  $x$  es una de las palabras en cuestión. Abra el archivo en Praat, mida los dos primeros formantes en el centro de la vocal inicial de cada palabra y ubique sus valores en el cuadro vocálico que se muestra a continuación.



5. Utilizando la grabación realizada en el punto anterior, mida el ancho de banda de los dos primeros formantes ( $B_1$ ,  $B_2$ ) de la vocal inicial de cada palabra. Realice una tabla y mida el valor promedio de  $B_n$ : ¿Qué diferencias encontró entre el valor individual de  $B_1$ ,  $B_2$  de cada vocal y el valor promedio?

6. Elija una grabación que contenga un buen número de vocales tónicas y átonas, mida su duración, elabore una tabla con los valores obtenidos y calcule su valor promedio. ¿Puede observar alguna tendencia en las mediciones? Discuta los resultados.

**Cuadro 2.1.** Atajos de teclado para usar en el Editor II

Opción en el menú	Teclas	Resultado
Get cursor	F6	Se obtiene el tiempo en que se ubica el cursor.
Get spectral power at cursor cross	F7	Praat muestra el valor de la potencia del espectro en $\text{Pa}^2/\text{Hz}$ .
View spectral slice	Ctrl + L	En una ventana emergente aparece un espectro FFT de banda ancha.
Get Pitch	F5	Praat muestra el valor medio de $F_0$ de una selección o del punto en que se ubica el cursor.
Get minimum pitch	Ctrl + F5	Praat muestra el valor mínimo de $F_0$ en el área seleccionada.
Get maximum pitch	Mayús + F5	Praat muestra el valor máximo de $F_0$ en el área seleccionada.
Get intensity	F8	Praat muestra el valor medio de la intensidad de una selección o del punto en que se ubica el cursor.

**Cuadro 2.2.** Atajos de teclado para usar en el Editor III

Opción en el menú	Teclas	Resultado
Get minimum intensity	Ctrl + F8	Praat muestra el valor mínimo de la intensidad en el área seleccionada.
Get maximum intensity	Mayús + F8	Praat muestra el valor máximo de la intensidad en el área seleccionada.
Get first formant	F1	Praat muestra el valor del primer formante en el punto en que se ubica el cursor.
Get second formant	F2	Praat muestra el valor del segundo formante en el punto en que se ubica el cursor.
Get third formant	F3	Praat muestra el valor del tercer formante en el punto en que se ubica el cursor.
Get fourth formant	F4	Praat muestra el valor del cuarto formante en el punto en que se ubica el cursor.
SoundEditor help	Ctrl + ?	Ayuda.
Search Manual	Ctrl + M	Buscar en la documentación oficial

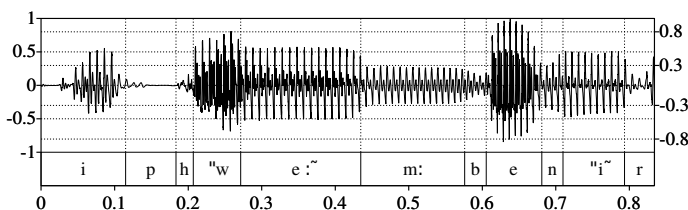


## 3 TRANSCRIPCIÓN Y ETIQUETADO

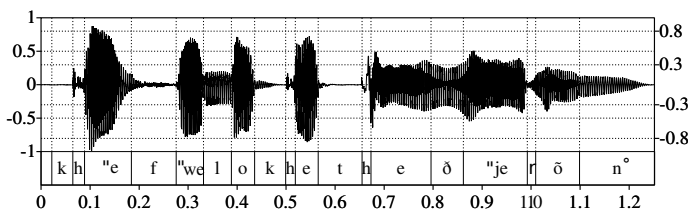
### 3.1 SEGMENTACIÓN

La *transcripción fonética* consiste en representar las propiedades articulatorias y acústicas de los sonidos del habla por medio de símbolos escritos. El *etiquetado* es una labor complementaria en la que se asignan marcas a ciertas variables para determinar su frecuencia de ocurrencia, la magnitud de sus correlaciones y, en general, para realizar mediciones y cálculos de manera automática. Más que una técnica propia de la profesión, la transcripción y el etiquetado de datos fonéticos se deben cultivar con la experiencia y, por tanto, no se aprenden directamente de un libro. Sin embargo, en este capítulo se ofrecen algunas explicaciones que ayudarán al lector a abordar este tipo de tareas.

Cuando analizamos una señal desde un punto de vista lingüístico, el primer problema que enfrentamos es descubrir la manera de dividir ondas continuas en categorías como las consonantes y vocales. El espectrograma no es de gran ayuda como primer acercamiento, pues la representación de los cambios acústicos no siempre muestra límites claros entre los sonidos de una lengua. La *forma de onda* u *oscilograma* es de mayor utilidad porque representa de manera simple las variaciones de *frecuencia* y *amplitud* de los *gestos articulatorios*. La figura 3.1 es la forma de onda de la interrogativa “¿pueden venir?” [i p<sup>h</sup>wē:m: be'nîr], y la figura 3.2 corresponde a la frase “¿qué fue lo que te dieron?” [k<sup>h</sup>e 'fwe lo k<sup>h</sup>e t<sup>h</sup>e 'ðjerõŋ] (hablantes bogotanos). Cada ejemplo está *segmentado* y *transcrito* fonéticamente para mostrar que los gestos articulatorios producen variaciones de amplitud y frecuencia con una *forma* y *duración* definida.



**Figura 3.1.** Segmentación de la onda correspondiente a la interrogativa *¿pueden venir?* [i 'p<sup>h</sup>wē:m: be'nir] teniendo en cuenta la *forma de onda, amplitud y frecuencia*



**Figura 3.2.** Segmentación de la onda correspondiente a la interrogativa *¿qué fue lo que te dieron?* [k<sup>h</sup>e 'fwe lo k<sup>h</sup>e t<sup>h</sup>e 'ðjerõŋ] teniendo en cuenta la *forma de onda, amplitud y frecuencia*

La primera distinción que puede hacerse es entre vocales *orales*, *vocales nasalizadas* y *laringalizadas*. En la figura 3.1 la vocal 'e' se realiza como un sonido oral [e] y como una vocal larga nasalizada en ['wē:]. La vocal oral tiene una amplitud de 1Pa, mientras que su contraparte nasalizada tiene una amplitud 0.5Pa debido a los efectos acústicos del tracto nasal y sus tejidos. La vocal laringalizada ['i~] tiene la misma amplitud que las nasalizadas (0,5Pa aprox.), sin embargo, sus pulsos glotales tienen menos frecuencia de ocurrencia que las vocales orales y nasalizadas. En general, cuando la frecuencia de ocurrencia de los pulsos glotales ( $F_0$ ) es alta, la forma de onda es compacta y oscura, mientras que cuando la frecuencia fundamental es baja los pulsos vocales

son pocos y pueden incluso contarse con una simple inspección p. ej., [i]. Con estas indicaciones también podemos saber que el hablante que produce la secuencia de la figura 3.1 es un hombre ( $F_0$  bajo), mientras que la persona que produce la secuencia de la figura 3.2 es una mujer ( $F_0$  alto).

Las consonantes nasales [m, n, ŋ] y la aproximante lateral [l] (figuras 3.1-3.2) tienen menos amplitud que las vocales nasalizadas (aprox. 0,3Pa) pero al igual que las vocales, se producen con pulsos glotales regulares i. e., son *sonoras*. Por la similitud de estos sonidos, en algunas teorías fonológicas las consonantes nasales, laterales y vibrantes se agrupan en una categoría llamada *sonantes*. Las consonantes oclusivas sordas [p, t, k] se producen con un cierre completo de los articuladores y, por tanto, la señal no muestra variaciones de amplitud o frecuencia. Como puede verse en los ejemplos, en español bogotano las oclusivas sordas son *aspiradas* [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>], es decir que después de la apertura de los articuladores hay una corriente de aire que pasa a través de la glotis.

Cuando trabajamos con señales capturadas en las condiciones adecuadas, la *forma de onda* ofrece parámetros claros para segmentar el habla en eventos individuales. Como primer acercamiento, podríamos decir que *la apertura de los articuladores es directamente proporcional a la amplitud de la señal*. Así, la llamada jerarquía universal de apertura articulatoria VOCALES > SONANTES > FRICATIVAS > OCLUSIVAS (Moreno, 1999, p.130) puede caracterizarse fonéticamente como variaciones sistemáticas en la *amplitud* y *frecuencia* de onda. Naturalmente, esta correlación no es absoluta; hay variabilidad inter e intralocutor, los gestos articulatorios se solapan, los sonidos resultantes se fusionan sintagmáticamente, etc. Para nuestros propósitos, lo importante es entender que las variaciones de *amplitud* y *frecuencia* asociadas al *modo de articulación* se identifican visualmente y sirven de base empírica para realizar una segmentación de la cadena de habla.



Praat ofrece diferentes opciones para transcribir y equiquetar fenómenos ‘segmentales’ y ‘suprasegmentales’. En lo que resta de este capítulo se estudiarán los procedimientos básicos para realizar estas tareas. Se hace énfasis en la caracterización de la curva melódica, pues el programa permite poner a prueba diferentes orientaciones teóricas y realizar alguna variedad de mediciones.

### 3.2 EL OBJETO **TextGrid**

Con el objeto **TextGrid** se transcriben y etiquetan señales de habla en Praat. Para crear un **TextGrid** desde la ventena de objetos elija las opciones **New** → **Create TextGrid**. En ese punto aparecerá una ventana con los siguientes campos:

1. **Start time(s)**: marque como tiempo de inicio 0.0.
2. **End time(s)**: el tiempo de finalización del **TextGrid** debe concordar con la duración del archivo de audio. Para conocer la duración exacta del sonido use el botón **Info** (parte inferior de la ventana de objetos) y en el texto emergente ubique la línea titulada **Total duration**.

Otra manera de realizar esta operación es crear el **TextGrid** sin especificar el tiempo de finalización. En ese caso seleccione el **TextGrid** junto con el archivo de audio y luego oprima **Scale times**. Debe tener presente que si no se realiza este paso el objeto **TextGrid** no quedará sincronizado con el audio y, por tanto, se perderá el trabajo realizado.

3. **All tier names**: asigne un nombre a cada uno de los niveles de la transcripción y sepárelos con un espacio p. ej., **palabra sílaba fono tono**, etc.
4. **Which of these are point tiers?**: Praat permite que los niveles de transcripción (**tiers**) estén constituidos por intervalos (p. ej., para delimitar la extensión de una palabra

y una sílaba) o marcas individuales ubicadas en un punto del tiempo (e.g. para identificar el centro de una vocal y extraer sus formantes de manera automática). En el primer caso se habla de un nivel de intervalos (**interval tier**) y en el segundo de un nivel de puntos (**point tier**). En esta casilla se debe especificar cuál de los niveles es un **point tier**; si desea usar intervalos solamente, deje el espacio en blanco.

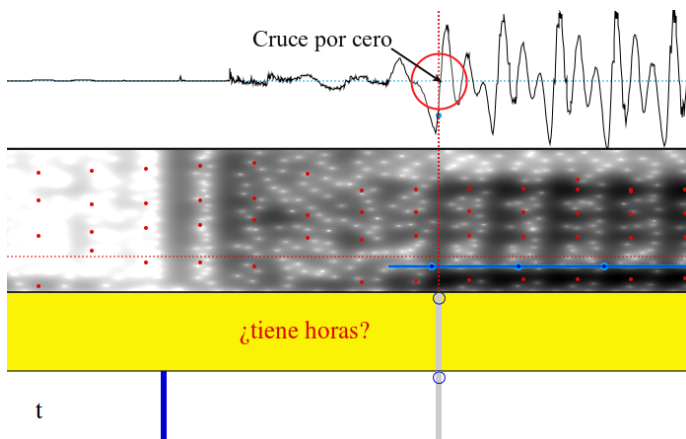
Una vez se han completado los campos anteriores, oprima **OK**, seleccione el **TextGrid** junto con el archivo de audio y oprima **View & Edit**. De esta manera podrá ver en el editor la forma de onda, un espectrograma de banda ancha y el **TextGrid** con los niveles que ha introducido previamente. En la figura 3.3, por ejemplo, introduje dos niveles de intervalos para transcribir un enunciado ortográficamente y segmentarlo en fonos.

Para establecer los límites (en el menú **Boundary**) de un intervalo debe hacer clic en la parte de la señal (espectrograma o forma de onda) en la que quiere establecer la marca. Como consecuencia aparecerá una línea vertical de color gris y con una circunferencia en la parte superior (nótese la ubicación del cursor en la figura 3.3); al hacer clic sobre esta, el límite se tornará azul, lo cual indica que el procedimiento se ha realizado correctamente.

Es recomendable que cada límite concuerde con un *cruce por cero* (i. e., aquella porción de la onda en que la presión no es positiva ni negativa) para evitar distorsiones en la reproducción y establecer marcas en puntos inadecuados (parte superior de la figura 3.3). Si desea ubicar de manera automática el cruce por cero más próximo a la selección, elija **Select → Move cursor to nearest zero crossing** o simplemente oprima **Ctrl + 0**. Como podrá notar, cuando se ubican los cruces por cero la *segmentación* es más precisa.

Si desea mover el límite (**Boundary**) tan solo debe seleccionarlo, mantener oprimido el clic derecho y desplazar el mouse hasta

el punto deseado. En el caso de que quiera eliminarlo puede ir al menú del Editor y elegir **Boundary** → **Remove** o simplemente usar las teclas **Alt + BackSpace**.



**Figura 3.3.** Anotación con un TextGrid y cruce por cero

Cuando realice la transcripción de un archivo de audio alineado con un TextGrid notará que aparecen nuevas opciones en el menú del Editor. Por ejemplo, en **File** ahora es posible dibujar un oscilograma y su transcripción con **Draw visible Sound and TextGrid**. De la misma manera, con **Draw visible pitch contour and TextGrid** podrá dibujar una curva de  $F_0$  asociada a un texto. Además, aparecen dos nuevas opciones en el menú (**Boundaries** y **Tier**) para editar los niveles de transcripción y los límites asignados. Al final de este capítulo se resumen los atajos de teclado para realizar operaciones rápidas en un TextGrid.

### 3.3 BÚSQUEDAS DESDE PRAAT OBJECTS

En el *menú dinámico* de la ventana de objetos hay opciones para extraer información del **TextGrid**. Estas opciones y sus comandos resultan útiles en la elaboración de **scripts**:

1. **List**: esta opción permite al usuario obtener una tabla **.txt** con información básica del **TextGrid**, como el tiempo de inicio de un intervalo (**tmin**), su nombre, el símbolo asignado (**text**) y el tiempo de finalización del intervalo (**tmax**).
2. **Down to table**: traslada los datos del **TextGrid** a una tabla de Praat. De esta manera se pueden aplicar procedimientos estadísticos sobre la información tabulada.
3. **Query**: en el menú dinámico elija **Query** → **Query time domain** para extraer información como el tiempo de inicio (**Get start time**), finalización (**Get end time**) y la duración total (**Get total duration**) de un **TextGrid**.

Para extraer el valor de los formantes se acostumbra a calcular el espectro en el centro de una vocal. Si desea hacerlo con un script puede obtener este valor calculando la diferencia entre el tiempo de finalización (**Get end time**) y el tiempo de inicio (**Get start time**) de un intervalo (**Interval tier**). Además, usted puede conocer el número de niveles de transcripción (**Get number of tiers**) o su nombre (**Get tier name ...**). Para conocer la manera en que está etiquetado un nivel acuda a **Query** → **is interval tier ...** Por ejemplo, si analizamos el primer nivel de un **TextGrid**, Praat arrojará un mensaje como:

1 (yes, tier 1 is an interval tier).

La opción **Query** permite extraer información de un nivel de intervalos o puntos. Aconsejo al lector explorar en el menú **Query interval tier**, **Query point tier** e incluirlas en las rutinas de sus scripts.

4. **Modify**: sirve para insertar límites (**Insert boundary...**), eliminar la marca izquierda (**Remove left boundary ...**), derecha (**Remove right boundary...**) o un límite ubicado en un punto específico (**Remove boundary at time...**). Si desea introducir texto (p. ej., un guión para identificar silencios) oprima **Modify** → **Set interval text** e introduzca los caracteres en el lugar que corresponda.
5. **Extract one tier...** es común introducir información en diferentes niveles de un **TextGrid**. Por ejemplo, es común iniciar con una transcripción ortográfica, luego segmentar en sílabas, fonemas, etc. Con esta opción se pueden extraer niveles y crear nuevos objetos para diferentes propósitos.
6. **Extract part**: puede acudir a esta alternativa cuando quiera extraer parte de un **TextGrid**. Para utilizarla necesita saber el tiempo de inicio y finalización del texto de interés.
7. **Analyse interval tier**: sirve para extraer información sobre el texto de un **TextGrid**. Por ejemplo, se puede buscar en un intervalo un carácter para identificar sus ocurrencias (**is equal to**), los símbolos diferentes (**is no equal to**), aquellos intervalos que contengan el carácter en cuestión (**contains**) o que inicien (**stars with**) y terminen con este (**ends with**).
8. **Analyse point tier**: ofrece las mismas alternativas que la opción anterior.
9. **Merge**: es posible mezclar dos o más objetos **Textgrid**; en ese caso Praat los ordenará uno debajo del otro. Esta

función resulta útil si usted cuenta con varios **TextGrid** de un mismo archivo con la misma duración.

10. **Concatenate**: es posible unir dos o más objetos **Textgrid**; en este caso Praat los ordenará uno después del otro. Esta función resulta útil si cuenta con varios **Textgrid** con la misma división i. e., el número de niveles de intervalos y puntos debe concordar.

### 3.4 EL ALFABETO FONÉTICO INTERNACIONAL

Para transcribir las vocales y consonantes de las lenguas del mundo actualmente se usa el *Alfabeto Fonético Internacional*. Se trata de un conjunto de símbolos y diacríticos acordados por los miembros de la [Asociación Fonética Internacional](#) [ði intə'næʃəŋəl fə'netɪk əsəʊsi'eɪʃn] en encuentros realizados desde 1886. El objetivo principal del alfabeto es representar aquellos sonidos que permiten distinguir las 'palabras' de una lengua i. e., sus fonemas (IPA, 1999; Ladefoged, 1990). Sin embargo, también se usa para representar variantes (fonos) relacionadas con el contexto fónico, la variedad dialectal o las características fónicas de un individuo.

Cuando representamos los fonemas de una lengua haciendo uso del menor número de símbolos, decimos que la transcripción es amplia (*broad transcription*), y cuando representamos los fonemas y fonos haciendo uso un número amplio de símbolos y diacríticos, decimos que la transcripción es estrecha (*narrow transcription*). Por ejemplo, el enunciado "¿en serio vas a venir?" tendría las siguientes representaciones para la variedad de español hablada en Bogotá (Colombia):

(1) /en 'serjo 'bas a be'nir/

(2) [ẽ.'se.rjo.'βa.sa.βe.'niɪ ↑]

En el primer ejemplo se usan letras del alfabeto latino para representar los sonidos distintivos del español (la única excepción es la vibrante simple *r* y el acento *´*). En el caso de la transcripción estrecha se usan símbolos especiales y diacríticos para representar variantes como la aproximante bilabial sonora [β̞], las vocales nasalizadas [ẽ̃ ï̃] y la aproximante alveolar [ɹ̠]. Además, se divide la secuencia en sílabas y se marca el ascenso final de  $F_0$  con el símbolo  $\uparrow$ . El tipo de transcripción que debemos usar depende del objetivo de la investigación. Por ejemplo, si el propósito es describir las vocales y consonantes de una lengua indígena, resultará conveniente usar una transcripción amplia; en el caso de que se quiera describir un dialecto del español o hacer un estudio sociolingüístico, una transcripción estrecha permitirá destacar las variantes que caracterizan la variedad lingüística de interés. Los profesionales interesados en patologías del habla pueden usar la extensión [extIPA SYMBOLS FOR DISORDERED SPEECH](#).

Es importante tener en cuenta que los fonemas y los fonos son conceptos que facilitan la comunicación de ideas sobre la lengua y el habla, pero no tienen ninguna realidad cognitiva como pretenden algunas teorías fonológicas (una discusión detallada se encuentra en Port, 2005, 2008). Para Öhman (2007, p.41), estas hipótesis están influenciadas por el uso del alfabeto, de manera que para muchos hablantes resulta difícil –y en algunos casos imposible– pensar en los sonidos del habla sin evocar la escritura. El autor opina:

In a manner of speaking the phonemic letter *cannot* exist without the corresponding speech sound being presented! The literate speaker has a command of many more concepts, and it is an urgent task of phonetics to work out the details of a native speakers' acoustic-phonetic conceptual repertoire.

A medida que el estudiante avance en la interpretación de la información acústica descubrirá que los hablantes utilizan un repertorio de sonidos más amplio de lo que supone la fonología segmental de su lengua. Por ejemplo, muchos de nosotros producimos la vocal final de los grupos fónicos con voz laringalizada (*creacky voice*). Claramente, este tipo de fonación no tiene importancia fonológica para una lengua como el español, pero es un recurso vocal que cualquier locutor puede usar para indicar el final de una preferencia, un estado anímico, entre otras cosas.

### 3.5 ALFABETO SAMPA

El alfabeto SAMPA ([Speech Assessment Methods Phonetic Alphabet](#)) es un sucedáneo del Alfabeto Fonético Internacional que está constituido por caracteres ASCII de 7 bits ubicados en el rango 33-127 (Wells, 1997). Este sistema fue creado en la década del ochenta por fonetistas de la comunidad europea para transcribir lenguas como el holandés, danés, italiano, alemán y el francés; luego el sistema fue ampliado para representar el noruego, sueco, griego y la lengua española (Llisterri y Mariño, 1993). La idea de fondo de SAMPA es realizar transcripciones que puedan ser procesadas por una computadora. En un principio el sistema tan solo permitía realizar transcripciones fonológicas, pero en la actualidad es posible utilizar extensiones para representar rasgos prosódicos ([SAMPROSA](#)) y realizaciones fonéticas ([X-SAMPA](#)). El alfabeto se utiliza en el campo de las tecnologías del habla para realizar corpus destinados a la *conversión de texto en habla* o el *reconocimiento automático*.

Es común ajustar las convenciones del alfabeto SAMPA para facilitar el análisis de un conjunto de variables o realizar scripts con un lenguaje de programación específico. Por ejemplo, en el Corpus de Kiel se añadieron etiquetas para estudiar fenómenos del habla espontánea como la glotalización, la laringalización (\$- q), la nasalización (\$- ~), la aspiración (\$- h), la elisión



de sonidos (\$ -), la presencia de correlatos acústicos de segmentos elididos (\$-MA) y etiquetas para las vocales de las palabras funcionales (a+) (Kohler, 1998). Por ejemplo, la transcripción estrecha de la frase *la niña modela en Roma* (hablante bogotano) con IPA es [lə 'niɲã mo'ela ɛn 'roma], con los caracteres X-SAMPA [l@ n"i~Ja~ mo"ela e\_K rr"om\_Ka\_K] y, finalmente, [l@+n\$"i~J\$a~ mo\$D-"ela \$e+-qnrr"o\$m-q\$a-q] es la adaptación SAMPA/KIEL. Las tres transcripciones son representaciones del mismo evento, sin embargo, la versión SAMPA/KIEL tiene una ventaja sobre las demás: los sonidos que se eliden se conservan en la transcripción (p. ej., [mo"ela] vs [mo\$D-"ela]), se introducen etiquetas para diferenciar el tipo de palabra (nótese que a los artículos y preposiciones les sigue el signo +) y fenómenos del habla espontánea como la glotalización, elisión de sonidos, fusión de palabras, etc. Con esto quiero llamar la atención sobre un aspecto importante del etiquetado de datos: *las variables consideradas en un estudio deben tener etiquetas que las identifiquen dentro del corpus*. Esto facilitará el tratamiento estadístico de las variables, descubrir correlaciones y la verificación empírica de la(s) hipótesis.

### 3.6 TRANSCRIPCIÓN Y ETIQUETADO DE LA ENTONACIÓN

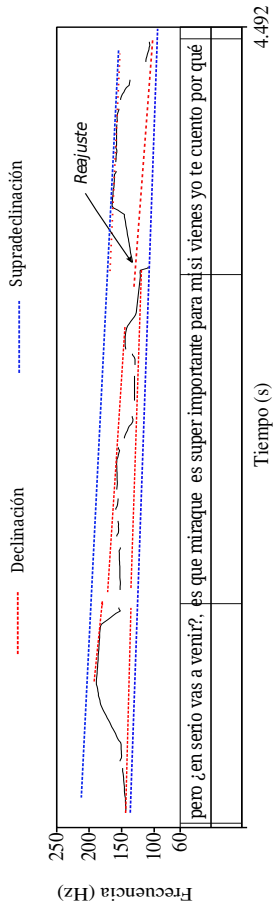
En la actualidad, uno de los temas que suscita mayor interés entre los lingüistas es la entonación. Cutler y Ladd (1983, p.5) diferencian entre enfoques descriptivos y enfoques experimentales. Los primeros caracterizan los contornos de  $F_0$  como secuencias de unidades discretas que se oponen paradigmáticamente (Pierrehumbert, 1980; Ladd, 1996, entre otros). En tanto, los enfoques experimentales relacionan de manera directa las características acústicas y perceptivas de  $F_0$  con funciones lingüísticas. Una clave para comprender el potencial de estos modelos de la entonación es entender cómo transcriben y etiquetan las variaciones de  $F_0$ .

En los apartados que siguen se expondrá cómo realizar el etiquetado fonético de la curva melódica y cómo hacer transcripciones fonológicas de la entonación con Praat. De paso, se hará una presentación sucinta de tres modelos que han sido utilizados para analizar el español: el sistema IPO, el Análisis Melódico del Habla y el sistema de notación Sp\_ToBI.

### 3.7 ETIQUETADO FONÉTICO DE LA CURVA MELÓDICA

El modelo IPO fue desarrollado en el *Institute for Perception Research* por Johan 't Hart, Antonie Cohen y René Collier (1990); unos años después fue aplicado a lenguas como el francés, inglés y alemán; su adaptación y aplicación al español se inició con el trabajo de Garrido (1996) y ha continuado con los trabajos del autor y sus colegas. Este modelo de base fonética se basa en el supuesto de que las curvas melódicas exhiben *formas* recurrentes que son independientes del hablante y del enunciado. Su objetivo principal es establecer un modelo predictivo de la forma de las curvas melódicas, determinar sus constituyentes básicos y, finalmente, establecer la función comunicativa de las configuraciones recurrentes (Garrido, 2003, p.97).

Las variaciones  $F_0$  durante la producción de uno o más enunciados se conoce como *curva melódica* (Estruch et al., 2007). Dentro del sistema IPO la curva se concibe como la superposición de *patrones globales* y *patrones locales*. Los patrones globales son la tendencia de las curvas melódicas a reducir su nivel tonal a lo largo del grupo entonativo (*declinación*) y a lo largo de grupos consecutivos (*supradclinación*). En la figura 3.4 se ilustran estas unidades en la frase “pero ¿en serio vas a venir?, es que mira que es super importante para mí, si vienes yo cuento por qué”, producida por un locutor bogotano. Nótese que el hablante produce tres unidades entonativas dentro de un solo patrón de supradclinación. Es decir, cada grupo entonativo tiene una forma inividual que se agrupa a su vez en una configuración global.

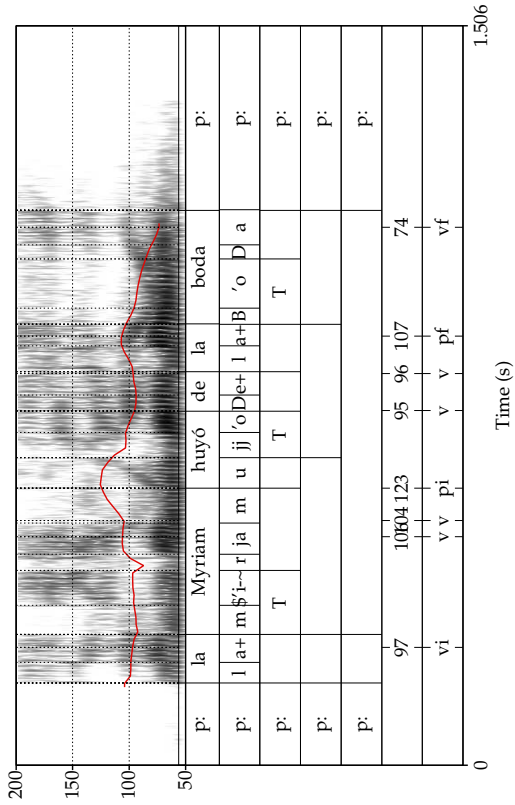


**Figura 3.4.** Unidades de análisis del sistema IPO. *Cláusula entonativa* constituida por tres *grupos entonativos* “pero ¿en serio vas a venir?”, es que mira que es super importante para mí, si vienes yo cuento por qué”, producida por un locutor bogotano

Los patrones locales son movimientos tonales que se dan dentro de una porción de vocal o dentro de una o más sílabas de un enunciado (Garrido, 2003). Estos movimientos de  $F_0$  se conocen como puntos de inflexión y se modelan como una línea recta. La unión de dos o más puntos de inflexión da lugar a *picos* y *valles*.

Para etiquetar manualmente un corpus en Praat haciendo uso de este sistema se deben seguir los siguientes pasos:

1. **Transcripción ortográfica y segmental:** como se ve en la figura 3.5, se inicia con una transcripción en palabras ortográficas y en sonidos usando niveles de intervalo (*interval tiers*). En este caso se utilizó el alfabeto SAM-PA para transcribir vocales y consonantes.
2. **Anotar las sílabas tónicas:** en un nuevo nivel se delimitan las sílabas tónicas y se marcan con una T.
3. **Delimitar los grupos acentuales:** el grupo acentual (también llamado *palabra prosódica* o *palabra fónica*) es la porción de enunciado que comprende el inicio de una sílaba tónica y la tónica siguiente (es el campo de desplazamiento de los picos tonales).
4. **Identificar los puntos de inflexión:** el penúltimo nivel ( $f_0$ ) se usa para etiquetar el valor absoluto en hercios de cada punto de inflexión. Para establecer estas marcas debe ubicar los momentos del tiempo en que hay movimientos relevantes de  $F_0$ . Dado que solo se marcan los movimientos importantes, se habla de una *estilización ancha* (Estruch et al., 2007).
5. **Identificar picos y valles:** en este nivel se especifica si el punto de inflexión marcado en el nivel anterior es un pico (P) o un valle (V).



**Figura 3.5.** Niveles de transcripción en el sistema IPO (véase el texto para mayor explicación)

Con la información anotada podrá realizar la medición del *rango* y la *declinación* de uno o más enunciados (Garrido et al., 1995). El rango es el valor máximo y mínimo de  $F_0$  y se calcula hallando la diferencia entre el pico máximo y el valle mínimo ( $p_{max}-v_{min}$ ). En tanto, la *declinación* se expresa en **sílabas/Hz** y se calcula estableciendo la diferencia entre el pico inicial y el pico final ( $p_i-p_f$ ), y entre valle inicial y valle final ( $v_i-v_f$ ) (Garrido et al., 1995).

### 3.8 NORMALIZACIÓN DE LA CURVA MELÓDICA

El Análisis melódico del habla (AMH) es un modelo fonético fonológico de la entonación en el que se analiza la curva de  $F_0$  independientemente de la estructura rítmica de la lengua. Según sus proponentes (Cantero y Font, 2009), el objetivo es brindar un conjunto de criterios de segmentación y procesamiento de los datos acústicos que faciliten la identificación de patrones fonológicos, estudiar la curva melódica en el habla espontánea, comparar resultados de múltiples informantes y aplicar los resultados dentro y fuera del aula. Se trata de un modelo *reversible*, pues el tratamiento de  $F_0$  permite sintetizar y resintetizar enunciados con información de un corpus.

Por otra parte, el AMH ofrece un procedimiento para *normalizar* las diferencias interindividuales relacionadas con la variedad de habla, el género, edad y otras características de los hablantes. La normalización consiste en analizar el contorno melódico como una sucesión de *valores relativos* y no solo como una sucesión de *valores absolutos* expresados en hercios. Este procedimiento resulta sumamente útil para identificar contornos entonativos que ocurren sistemáticamente en una lengua y puede ser aplicado para el estudio fonológico de la entonación. El modelo se basa en la siguiente jerarquía fónica:

1. **Grupo fónico:** es la unidad fonética que organiza y estructura el discurso oral; se le reconoce porque está delimitado por una inflexión tonal que funciona como núcleo (tonema).
2. **Grupo rítmico:** así se denomina ‘al conjunto de sonidos agrupados en torno a un acento’ (Cantero, 2002, p.52). Esta unidad comprende la porción de habla que se encuentra entre dos sílabas acentuadas; en otros modelos se denomina grupo acentual o palabra fónica.
3. **Segmento tonal:** es ‘cada uno de los estadios tonales más o menos estables y claramente perceptibles’, que suelen coincidir con una mora [...]. Una vocal suele durar una mora’. (Cantero, 2002, p.89).

En la figura 3.6 se muestra el etiquetado de la interrogativa “¿puedo pasar?”, realizada según los principios del AMH. En este caso utilicé un nivel de intervalo para segmentar las sílabas y tres niveles de punto para etiquetar el valor en hercios (Hz), calcular el valor relativo de la sucesión de tonos (**Porcentaje**) y realizar una curva estandarizada (**Curva estándar**). Como se ve, el grupo fónico tiene tres sílabas [ˈpoː .pa.ˈʃaːɪ], las cuales están constituidas por *segmentos tonales* (vocales tónicas o átonas) que experimentan variaciones de  $F_0$  y consonantes que resultan marginales para el estudio del contorno. La vocal tónica portadora de una inflexión tonal (p. ej., [ˈʃaːɪ]) constituye el núcleo del grupo fónico y generalmente tiene la duración de dos segmentos tonales (Cantero y Font, 2009, p.23). Para etiquetar los valores de  $F_0$  relevantes, los cambios relativos y la curva estándar se deben seguir los siguientes pasos:

1. **Ubicar puntos de inflexión:** son los puntos en que ocurren variaciones relevantes de  $F_0$ . Según Cantero y Font (2009) el valor en Hz puede calcularse en un punto estable

de la vocal (*vowel target*), considerarse el valor central (*moda o media*) o tomar los valores extremos de la inflexión. Lo fundamental de establecer cada etiqueta es recuperar el contorno melódico sin alterar, con las marcas asignadas, el análisis de las características acústicas. En el grupo fónico de la figura 3.6 los valores absolutos de los puntos de inflexión son: 236 205 256 247 345 Hz.

2. **Calcular los valores relativos:** en este paso la diferencia en hercios entre los puntos de inflexión se transforma a porcentajes para obtener los valores relativos de los cambios del contorno. Para realizar esta conversión haga un nivel de puntos llamado ‘Porcentaje’, y al primer punto de inflexión ( $p_1$ ) asigne el valor de 100 %. Para establecer el valor relativo del segundo punto de inflexión ( $p_2$ ) calcule la diferencia en hercios entre  $p_2$  y  $p_1$  ( $205-236 = -31\text{Hz}$ ); si el resultado es negativo significa que hay un descenso de  $F_0$ , y si el resultado es positivo significa que hay un ascenso. Para conocer el valor relativo de la diferencia, se realiza una regla de tres simple:

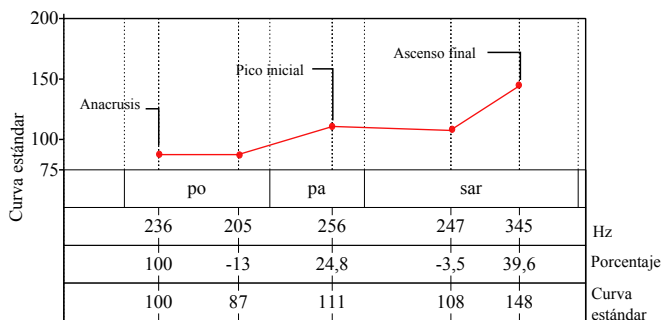
$$236 \text{ Hz} \rightarrow 100 \%$$

$$-31 \text{ Hz} \rightarrow x$$

El procedimiento puede utilizarse para hallar los demás valores. En el grupo fónico de la figura 3.6 los valores relativos resultantes fueron: 100 %, -13.1 %, 24.8 %, -3.5 %, 39.6 %.

3. **Establecer la curva estándar:** para representar los cambios del contorno se convierten los porcentajes a valores estándar. Para este fin se asigna el valor de 100 al primer pico y se suman o restan los cambios relativos e.g 100 -13,1 % + 24,8 % -3,5 % +39,6 %. Para el contorno de la figura 3.6 los resultados obtenidos son: 100 87 111 108 148.





**Figura 3.6.** Transcripción y etiquetado de un grupo fónico con el modelo de Análisis Melódico del Habla (Cantero, 2002)

Los datos correspondientes a la curva estándar pueden introducirse en un programa de análisis estadístico para obtener un diagrama de línea (figura 3.6). Por ejemplo, introduzca los valores estándar en una fila de Excel, seleccione las casillas y elabore el diagrama que representa el patrón melódico.

### 3.9 TRANSCRIPCIÓN FONOLÓGICA

El sistema ToBi está basado en el modelo métrico autosegmental de la entonación. Fue diseñado para el inglés (Pierrehumbert, 1980), pero luego fue adaptado a otras lenguas como el francés, italiano, holandés, alemán, griego, portugués, catalán, etc. La adecuación al español (Sp\_ToBI) se inició con los trabajos de Sosa (1999), Beckman et al. (2002), y ha sido continuada con los trabajos realizados por Face (2003), Face y Prieto (2007), Estebas y Prieto (2008), Prieto y Roseano (2010), entre otros.

El sistema ToBI describe la entonación en términos de una sucesión de tonos altos (H), bajos (L) y combinaciones de estos. Dado que para la teoría autosegmental las variaciones tonales se superponen al acento, se suelen marcar con asterisco (\*) los tonos alineados a las sílabas acentuadas p. ej.,  $L^*+H$  o  $L+H^*$ . Para transcribir el movimiento tonal que ocurre al final de una frase se utiliza el signo % p. ej., L %, H %, M %. En tanto, las junturas que ocurren en las frases intermedias se transcriben -H, -M, -HH, etc.

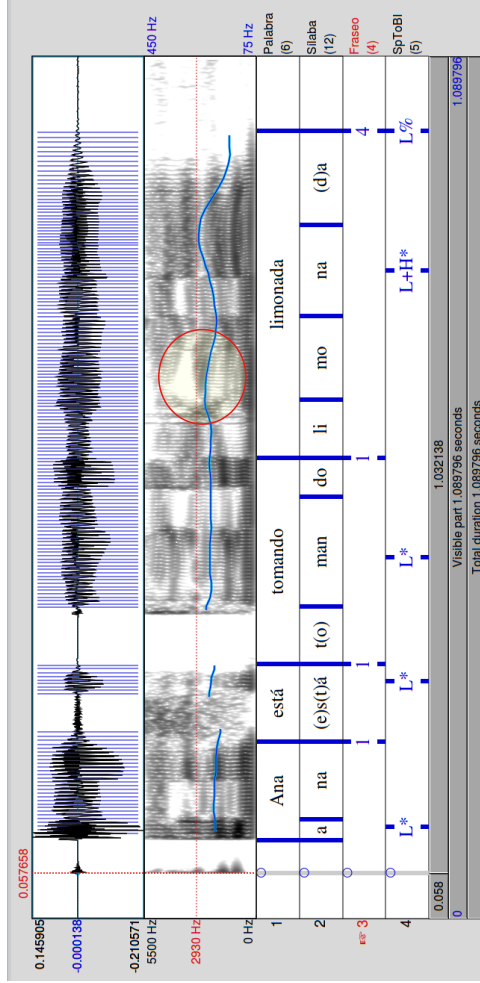
Las categorías fonológicas principales del modelo son los *acentos tonales* (i. e., las variaciones de  $F_0$  asociadas a las sílabas acentuadas) y los tonos de frontera (i. e., los movimientos tonales con los que finaliza una frase entonativa). De acuerdo con Estebas y Prieto (2008, p.2) los niveles considerados en el análisis son:

1. **Nivel ortográfico:** se representan los enunciados utilizando la ortografía de la lengua y se divide la secuencia en palabras y/o sílabas. En Praat puede usar niveles de intervalo para anotar este nivel.
2. **Nivel tonal:** se representa el contorno melódico con secuencias de acentos monotonaes o bitonaes. En Praat puede usar niveles de intervalo o de puntos para anotar este nivel.
3. **Nivel de separación prosódica:** en este nivel se marca el límite de palabras prosódicas, frases intermedias y frases entonativas con los números 0 1 2 3 4. En Praat puede usar niveles de punto para anotar este nivel.
4. **Nivel misceláneo:** sirve para etiquetar todos aquellos elementos extralingüísticos que dificultan el análisis o interpretación del corpus e.g. risas, cualidad de la voz, dubitaciones, etc. Si para su investigación es útil incluir esta información, un nivel de intervalos resultará conveniente.

En la figura 3.7 se muestra la transcripción de la declarativa “Ana está tomando limonada”, producida por una mujer hablante de español bogotano. El **TextGrid** fue creado con dos niveles de intervalo (**Palabra Sílabas**) y dos niveles de punto (**Fraseo Acento tonal**). En el primero se segmentó la señal en palabras ortográficas, en el segundo se delimitaron las sílabas y se indicó por medio de paréntesis los sonidos elididos. En el tercer nivel se estableció el límite de las palabras prosódicas (número 1) y con 4 el límite de la frase entonativa (el número 2 se utiliza para límites audibles pero no evidentes en el espectrograma, y el número 3 se utiliza para marcar el límite entre frases intermedias). Por último, se identificaron los acentos tonales asociados a las sílabas acentuadas y a los tonos de frontera.

El sistema Sp\_ToBI presenta algunos problemas de implementación. Por ejemplo, si observamos con detenimiento la figura 3.7, es evidente que hay dos picos de  $F_0$  en la palabra *limonada*. Dado que los movimientos tonales dependen del acento, el primer pico se descarta del análisis. En el caso de que descubriéramos que el pico está asociado a la sílaba acentuada de la palabra anterior (i.e. *tomando*), el sistema no ofrece convención alguna para indicar que se ha desplazado dos sílabas (existe **L+H\*** y **L+>H\*** pero no una tercera opción). Pero si existiera tal convención, valdría la pena preguntarse si el problema se debe a la falta de consenso sobre un conjunto de símbolos o un problema derivado de la teoría. En todo caso, es preciso aclarar que el sistema está en constante desarrollo, de manera que es probable que se mejore la notación y se solucionen los problemas teóricos que plantea su aplicación.

Una de las principales ventajas de la notación fonológica es su creciente uso para describir lenguas y dialectos. Esto facilita enormemente la comparación y la generalización de los resultados. Los métodos fonéticos de anotación están menos difundidos y, por tanto, resulta más difícil decidir qué rasgos son particulares de un dialecto.



## 3.10 BIBLIOGRAFÍA RECOMENDADA

Cantero, Francisco José. (2002). *Teoría y análisis de la entonación*. Barcelona: Universidad de Barcelona. La primera parte es un estado de la cuestión que resume las tradiciones y modelos en el estudio de la entonación. La segunda parte es una discusión entorno al *acento* y su relación con el material segmental y con la curva melódica. En la tercera y última parte, el autor presenta el fundamento teórico y empírico del *Análisis melódico del habla*.

Estebas, Eva y Prieto, Pilar. (2008). La notación prosódica en español. Una revisión del Sp\_ToBI. *Estudios de Fonética Experimental*, XVII, p. 263-283. Este artículo es un tutorial breve y concreto del sistema Sp\_ToBI. En este documento el lector encontrará las pautas para la asociación de acentos tonales con movimientos de  $F_0$ . Recuperado en: <http://stel.ub.edu/labfon/sites/default/files/XVII-15.pdf>.

Estruch, Mónica; Garrido, Juan María; Llisterri, Joaquim, y Riera, Montserrat. (2007). Técnicas y procedimientos para la representación de las curvas melódicas. *RLA. Revista de lingüística teórica y aplicada*, 45(2), 59-87. En este artículo se discuten los métodos para *calcular, estilizar y anotar* fonéticamente una curva melódica. Esta lectura es recomendada para aquellas personas interesadas en el estudio acústico y perceptivo de  $F_0$ . Recuperado en: [http://liceu.uab.cat/~joaquim/publicacions/Sevilla\\_99.pdf](http://liceu.uab.cat/~joaquim/publicacions/Sevilla_99.pdf).

Hidalgo, Antonio. (2006). *Aspectos de la entonación española: viejos y nuevos enfoques*. Madrid: Arco Libros. Es una introducción breve y amable al estudio de la entonación. La primera parte define la entonación y sus diferentes ámbitos, luego se ocupa de exponer los diferentes tipos de entonación lingüística y cierra con una sección dedicada a explicar los modelos de la entonación abordados aquí.

Prieto, Pilar (ed.). (2003). *Teorías de la entonación*. Barcelona: Ariel. Este libro presenta en ocho capítulos diferentes tendencias en el estudio de la entonación. El lector encontrará capítulos dedicados al Modelo Métrico-Autosegmental, al sistema IPO, a la escuela británica, al modelo Aix-en-Provence y, finalmente, un capítulo sobre la aplicación de los modelos de la entonación al campo de las tecnologías del habla. Este trabajo es recomendado para aquellas personas que quieren tener una visión de conjunto.

Prieto, Pilar y Roseano, Paolo (coords.). (2010). *Transcription of Intonation of the Spanish Language*. Lincom Europa: München. Este trabajo es un estudio sistemático de la fonología de la entonación del español hablado en España (Cantabria, Castilla, Islas Canarias) y América (Chile, Ecuador, Venezuela, Argentina, México, República Dominicana) utilizando el sistema Sp\_ToBI.

Sosa, Juan Manuel. (1999). *La entonación del español. Su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra. Esta obra es una de las primeras aplicaciones del sistema ToBi al análisis del español. En la primera parte se hace una presentación general del estudio de la entonación española, se discuten las categorías y se presenta la teoría de Pierrehumbert (1983). El segundo capítulo hace una descripción de la fonología de la entonación del español basada en un amplio corpus de español americano y peninsular. El libro cierra con un capítulo dedicado a la dialectología de la entonación del español.

### 3.11 RECURSOS RECOMENDADOS

En Internet se encuentran disponibles varios manuales de Praat y conjuntos de **Scripts** creados por diferentes investigadores. Recomendando especialmente los siguientes:

Correa, José Alejandro. (2013). *Díptico de alfabetos fonéticos*. Este documento presenta las equivalencias entre el Alfabeto Fonético Internacional, X- SAMPA y los comandos para usar las fuentes fonéticas en Praat. Adicionalmente, contiene un resumen de los sonidos del español con una tabla de equivalencias entre el Alfabeto Fonético Internacional (IPA) y el Alfabeto de la Revista de Filología Española (RFE).

ENLACE: <http://www.bibliodigitalcaroycuervo.gov.co/962/>

Garrido, Juan María. (2013). *SegProso*: se trata de un conjunto de **Scripts** para hacer etiquetado automático de la curva melódica según la adaptación del autor del Modelo IPO. En la página del proyecto encontrará un ejecutable con los archivos para Praat. ENLACE: <http://www.upf.edu/glicom/en/recursos/eines/segproso.html>

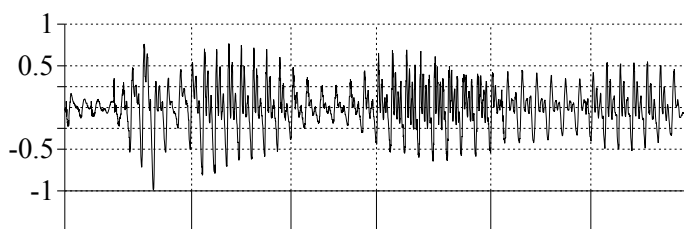
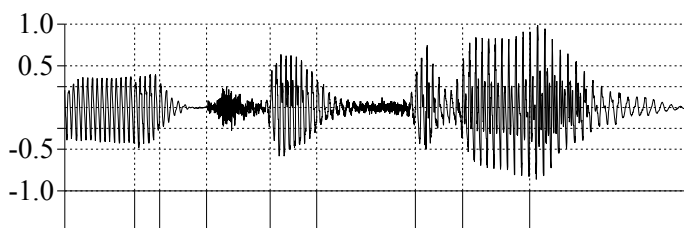
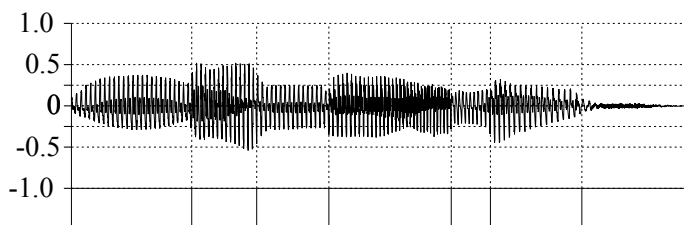
Goldman, Jean Philippe. (2011). *EasyAlign: an automatic phonetic alignment tool under Praat*. Es un complemento (*plugin*) para realizar transcripciones semiautomáticas del inglés, francés, español y taiwanés. El resultado es un **TextGrid** con los niveles *ortográfico, fonológico, palabras, sílabas y fonos*. En la página del proyecto encontrará las instrucciones de descarga y documentación en francés, inglés y español. ENLACE: <http://latlcui.unige.ch/phonetique/easyalign.php>

Lenne, Mieta. *The Speech Corpus Toolkit for Praat*. En esta página la autora pone a disposición un conjunto de **scripts** para realizar tareas como medir la duración de los sonidos etiquetados en un **TextGrid**, alinear límites (**Boundaries**), reemplazar texto, etc. Para comprender el uso de los recursos es necesario que el lector tenga dominio de las funciones básicas de Praat.

ENLACE: <http://www.helsinki.fi/~lenne/praat-scripts/>

## 3.12 EJERCICIOS SUGERIDOS

1. Los siguientes oscilogramas o formas de onda están etiquetadas segmentalmente. Identifique cuál corresponde a las palabras *limones*, *regalo* y a la expresión *mucho frío*. ¿Qué relación encuentra entre la amplitud de onda y el modo de articulación? Argumente su respuesta con base en los ejemplos.



2. Grabe una declarativa simple (p. ej., “Mi madre ama la moda”) y una interrogativa categórica (p. ej., *¿Viene a bailar?*) con



sonidos sonoros. Luego elabore un **TextGrid** con dos niveles de intervalo. En el primero segmente la oración por palabras y en el segundo divida la secuencia en segmentos y transcriba con el alfabeto SAMPA. Escriba cuáles fueron sus principales dificultades para realizar la tarea; en lo posible discuta los resultados con sus compañeros de estudio.

3. Añada un nivel de intervalos (**tier**) a la anotación de las frases del punto anterior. Delimite por sílabas cada oración y asigne un acento tonal del sistema Sp\_ToBI a las sílabas tónicas (i. e., L\*, H\*, L+H\*, L+H\*, H+L\*, L+>H\*, etc.). Contraste los resultados con la propuesta de Estebas (2008) antes de llegar a una transcripción definitiva.

4. Como se vio, el *rango* de  $F_0$  se calcula hallando la diferencia entre el pico máximo y el valle más bajo ( $p_{max}-v_{min}$ ). En tanto, la *declinación* se expresa en *sílabas/Hz*, y se define como la diferencia entre el pico inicial y el pico final ( $p_i-p_f$ ), y entre el valle inicial y el valle final ( $v_i-v_f$ ). Determine el rango y la declinación de la oración declarativa que viene trabajando. Para realizar la tarea, añada un nivel de puntos al **TextGrid**, y ubique los puntos de inflexión de la curva melódica como se hace en el sistema IPO (figura 3.5).

5. Una oración tiene puntos de inflexión correspondientes a 99 97 98 114 92 108 96 96 138 hercios (Hz). Calcule la curva estándar según el planteamiento del Análisis Melódico del Habla. ¿De qué tipo de oración se trata?

6. Instale la última versión de Praat y habilite *EasyAlign* y *Seg-Proso*. Realice una transcripción automática de las oraciones analizadas y compare los resultados con la anotación realizada previamente. Discuta los resultados con sus compañeros de estudio.

**Cuadro 3.1.** Atajos de teclado para usar en el Editor IV

Opción en el menú	Teclas	Resultado
Find ...	Ctrl + F	Buscar información en un TextGrid
Find again	Ctrl + G	Buscar nuevamente información en un TextGrid
Select previous tier	Alt + Up	Seleccionar el nivel previo
Select next tier	Alt + Down	Seleccionar el siguiente nivel
Select previous interval	Alt + Left	Seleccionar el intervalo previo
Select next interval	Alt + Right	Seleccionar el próximo intervalo
Extend -select left	Shift + Alt + Left	Seleccionar los intervalos hacia la izquierda
Extend -select right	Shift + Alt + Right	Seleccionar los intervalos hacia la derecha

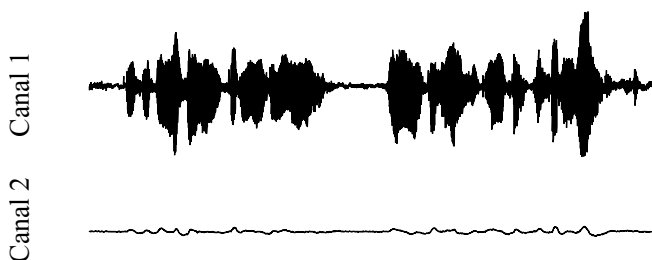
**Cuadro 3.2.** Atajos de teclado para usar en el Editor V

Opción en el menú	Teclas	Resultado
Add on tier 1	Ctrl + F1	Añadir un nivel sobre el tier 1
Add on tier 2	Ctrl + F2	Añadir un nivel sobre el tier 2
Add on tier 3	Ctrl + F3	Añadir un nivel sobre el tier 3
Add on tier 4	Ctrl + F4	Añadir un nivel sobre el tier 4
Add on tier 5	Ctrl + F5	Añadir un nivel sobre el tier 5
Add on tier 6	Ctrl + F6	Añadir un nivel sobre el tier 6
Add on tier 7	Ctrl + F7	Añadir un nivel sobre el tier 7
Add on tier 8	Ctrl + F8	Añadir un nivel sobre el tier 8
Add on all tiers	Ctrl + F9	Añadir un nivel sobre cada tier
Remove	Alt + BackSpace	Eliminar nivel

## 4 MANIPULACIÓN DE LA SEÑAL

### 4.1 EDICIÓN DE UNA SEÑAL DIGITAL

Las muestras de habla deben capturarse en *mono*. Sin embargo, esta opción no siempre funciona en las grabadoras digitales (figura 4.1) y, por tanto, es necesario editar antes de un análisis. Para convertir un archivo *stereo* a *mono* use el botón **Convert** ubicado en la ventana de objetos. Una alternativa es elegir **Convert** → **Convert to mono**, en cuyo caso se unificarán ambos canales. Otra alternativa es extraer el mejor canal con **Convert** → **Extract one channel...** Por ejemplo, en una señal como la que se muestra en la figura 4.1 es preferible seleccionar el canal 1 y no promediar toda la información.



**Figura 4.1.** Muestra en *mono* con un canal degradado

De otra parte, es necesario normalizar la escala de amplitud. Praat asume que la forma de onda tiene valores entre 1.0 y -1.0, sin embargo, el valor real depende de las condiciones originales de grabación. Para ajustar la amplitud debe elegir **Modify** → **Scale peak...** En la ventana emergente debe elegir la opción

estándar (0.99) y oprimir OK. De esta manera, si la amplitud es baja, Praat multiplicará los valores, y si la amplitud es alta, los dividirá. En la figura 4.2 se ve una forma de onda antes y después de la normalización; el efecto es auditivo, así que con esta opción también se mejorara la calidad de sonidos poco audibles. *Si no se realiza este procedimiento, no podrá asignar adecuadamente los rótulos cuando elabore una figura de la forma de onda.*



**Figura 4.2.** Efecto de la opción Scale peak...

## 4.2 REMUESTREO

Luego de editar la señal, es necesario ajustar la *frecuencia de muestreo* al tipo de análisis. Generalmente se usa una tasa de 44.100Hz, lo cual significa que la frecuencia más alta en la señal es de 22.050Hz, es decir, la mitad de la tasa muestreo o *frecuencia nyquist*. Así, si desea analizar vocales en un rango entre 100 y 5500 Hz, deberá *remuestrear* la señal a 11.000Hz. Por ejemplo, para analizar formantes utilizando el método To LPC (burg), necesita remuestrear a esta tasa para que Praat calcule correctamente el espectro. De lo contrario, tendrá una representación

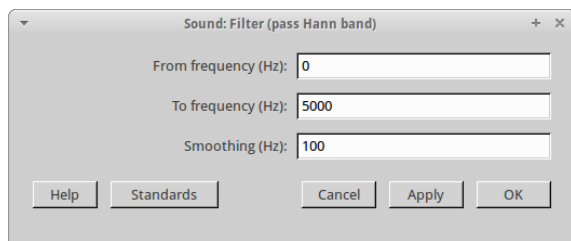
sin picos ni variaciones en el Nivel de intensidad. Para remuestrear una señal en Praat solo debe oprimir el botón **Convert** → **Resample...** ubicado en la ventana de objetos. En la ventana emergente debe ingresar la nueva tasa y oprimir OK.

### 4.3 FILTROS

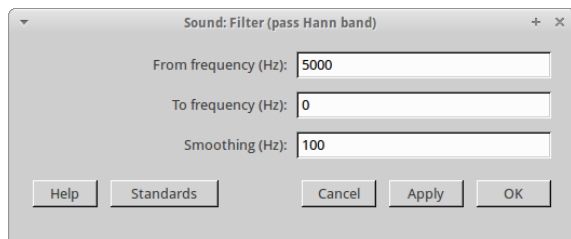
Los filtros acústicos han sido clave para la creación de técnicas de procesamiento de señales y para comprender el habla en general. De hecho, la *Teoría acústica de la producción del habla* (Chiba y Kajiyama, 1958; Fant, 1970) considera que el tracto vocal humano filtra las ondas cuasiperiódicas producidas por los pliegues vocales: las frecuencias que coinciden con las resonancias del tracto vocal se destacan (formantes) y las frecuencias restantes se atenúan. En la telefonía, por ejemplo, la señal está filtrada a 4.000Hz (estándar PCM) para evitar costos en la transmisión de los datos y para que el oyente pueda decodificar la información de un mensaje.

En acústica se trabaja con filtros *paso bajo*, filtros *paso-alto* y filtros *paso-banda*. En Praat pueden implementarse de la siguiente manera:

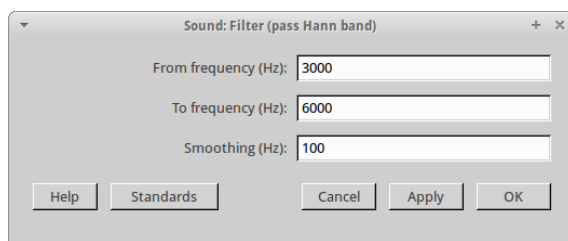
1. **Filtro paso-bajo:** admite las frecuencias que están por debajo de un valor determinado y elimina las restantes. Para implementarlo en Praat oprima **Filter-** → **Filter (Pass Hann band)**... En la ventana emergente la frecuencia inicial es cero (**From frequency (Hz): 0**) y el usuario debe asignar un valor máximo (**To frequency (Hz): valor máximo**). En el ejemplo mostrado a continuación, el filtro admite frecuencias hasta 5.000Hz. Adicionalmente, se hace un alisado de 100Hz (**Smoothing (Hz): 100Hz**) para evitar un corte drástico de las frecuencias.



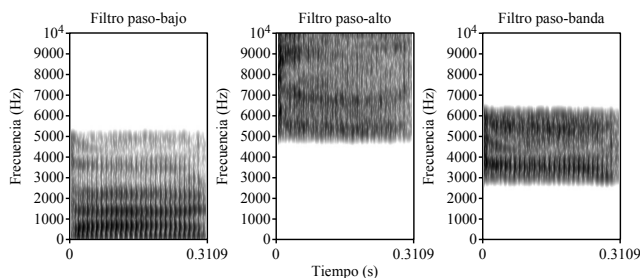
2. **Filtro paso-alto:** admite las frecuencias que están por encima de un valor determinado y elimina las restantes. Para usarlo en Praat oprima **Filter- → Filter (Pass Hann band)**... En la ventana emergente la frecuencia inicial es un valor mínimo asignado por el usuario (**From frequency (Hz): valor mínimo**) y el valor máximo es cero (**To frequency (Hz): 0**). En el ejemplo mostrado a continuación, el filtro admite frecuencias desde 5.000Hz.



3. **Filtro paso-banda:** admite las frecuencias que se ubican dentro de un rango determinado de frecuencias y elimina las restantes. Para hacerlo en Praat oprima el botón **Filter- → Filter (Pass Hann band)**... Luego asigne un valor mínimo (**From frequency (Hz): valor mínimo**) y un valor máximo (**To frequency (Hz): valor máximo**). En el ejemplo mostrado a continuación, el filtro admite frecuencias entre 3.000Hz y 6.000 Hz.



Como señala (Johnson, 2006), los filtros paso-banda pueden verse como la combinación de un filtro paso-bajo y un filtro paso-alto que tienen la misma frecuencia de corte. Así, el filtro paso-banda de la figura 4.3 (3.000 Hz-6.000Hz) también puede realizarse aplicando un filtro paso-bajo entre 0-6.000Hz, y un filtro paso alto desde 3000Hz hasta 6.000Hz. Se dice que las frecuencias que admite un filtro de este tipo (i. e., *su ancho de banda*) son aquellas que tienen el 70.7 % de la amplitud de la frecuencia central (Ladefoged, 1996). En nuestro ejemplo el rango de frecuencias está entre 3.000Hz y 6.000Hz y, por tanto, el ancho de banda es de 3.000Hz.



**Figura 4.3.** Filtros *paso-bajo*, *paso-alto* y *paso-banda*

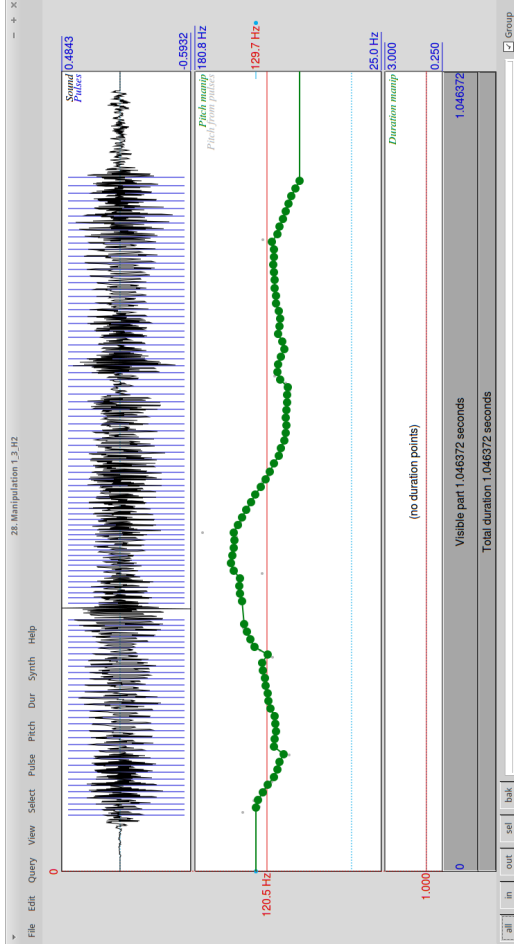


#### 4.4 TO MANIPULATION

Praat ofrece al usuario opciones para manipular rasgos prosódicos como la duración y el tono. Para modificar alguno de estos parámetros debe tener un archivo de audio en la ventana de objetos, oprimir **Manipulate** → **To manipulation...** y tendrá un objeto **Manipulation** en la ventana principal. Para visuar este fichero oprima **View & Edit** y tendrá acceso al Editor de manipulación (**ManipulationEditor**); a primera vista es semejante al editor general pero, a diferencia de este, no muestra el espectrograma y carece de opciones para realizar mediciones acústicas. Este editor contiene los siguientes elementos:

1. **Oscilograma:** la forma de onda con una delimitación de los pulsos glotales tomando como referencia el pico de máxima amplitud (parte superior de la figura 4.4). Puede añadir pulsos eligiendo en el menú **Pulse** → **Add pulse at cursor** o simplemente oprimiendo **Ctrl + P**.
2. **Nivel tonal:** *PitchTier* es una copia de la curva de  $F_0$  en la cual cada pulso glotal corresponde a un punto de inflexión (panel intermedio de la figura 4.4). Sin embargo, los puntos de inflexión se manipulan independientemente.
3. **Nivel de duración:** *DurationTier* sirve para manipular la duración de los sonidos reduciendo o aumentando la distancia entre dos puntos (recuadro inferior de la figura 4.4).

La manipulación controlada del tono y la duración le resultará útil para estilizar y resintetizar curvas melódicas, crear estímulos y evaluar sus efectos perceptivos, etc. Sin embargo, las *muestras manipuladas no deben usarse para extraer generalizaciones sobre las características prosódicas de un hablante, una variedad o una lengua.*



**Figura 4.4.** Editor de manipulación (ManipulationEditor): Forma de onda y pulsos glosales (parte superior), la curva de inflexión (panel intermedio) y nivel para manipular la duración (recuadro inferior)

#### 4.5 MANIPULACIÓN DEL TONO

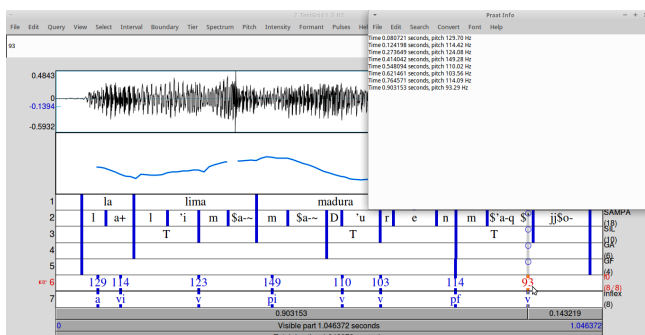
Cuando se abre el Editor de manipulación (Panel intermedio, figura 4.4), las variaciones de  $F_0$  se representan como una sucesión de puntos. Si selecciona uno o varios de estos podrá modificar su ubicación en el tiempo (desplazamientos horizontales) o su valor en hercios (movimientos verticales). En la parte superior del menú, la opción **Pitch** ofrece al usuario alternativas para modificar puntos de inflexión individuales o la configuración global del enunciado:

1. **Add pitch point at cursor**: añade un punto a la curva.
2. **Add pitch point at**: añade un punto de inflexión en un tiempo determinado.
3. **Set pitch range**: sirve para especificar el valor máximo que puede alcanzar  $F_0$ .
4. **Set pitch units**: la escala para el tono es en hercios o semitonos.
5. **Set pitch dragging strategy**: con esta opción se restringe la dirección de movimiento de los puntos de inflexión.
6. **Shift pitch frequencies**: con esta opción podemos realizar movimientos positivos o negativos en la totalidad del contorno.
7. **Multiply pitch frequencies**: multiplica la totalidad de la onda por un factor cuyo valor estándar es 1.2.
8. **Stylize Pitch**: reduce el número de puntos que conforman la curva de  $F_0$ .
9. **Stylize Pitch (2st)**: reduce a un mínimo de puntos de inflexión la curva de  $F_0$ .

#### 4.6 ESTILIZACIÓN DE LA CURVA MELÓDICA

Como se dijo en el capítulo anterior, las curvas melódicas se modelan fonéticamente con puntos de inflexión que forman picos y valles. Este principio permite simplificar las variaciones de  $F_0$  a un conjunto de movimientos relevantes o, en otras palabras, realizar una *estilización*. La *curva estilizada* resultante debe contener pocos puntos y ser perceptivamente igual a la original (Estruch et al., 2007). Veamos cómo estilizar la declarativa que se muestra en la figura 4.4.

1. El primer paso es tener la oración con la curva melódica cuidadosamente etiquetada en puntos de inflexión (véase el capítulo anterior).
2. Segundo, se debe recuperar el valor en hercios de cada punto y su ubicación temporal. Para hacer esto debe seleccionar cada punto en el **TextGrid** y oprimir la tecla **F12**. De esta manera tendrá una ventana de diálogo con la información requerida:



3. Tercero, debe crear un nivel tonal para sintetizar la curva melódica. Para ello elija en la ventana de objetos **New** →

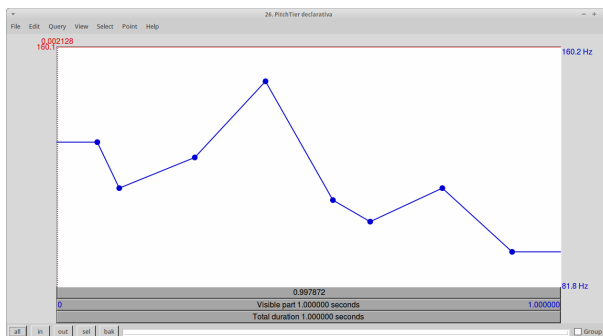
**Tiers** → **Create PitchTier**.... Un nivel tonal de un segundo de duración (1s) es suficiente para esta declarativa. Para ver el nivel tonal seleccione el nuevo objeto **PitchTier** y oprima **View & Edit**.

4. Cuarto, introduzca el valor en hercios de cada punto de inflexión y su ubicación temporal (paso 1). Puede hacerlo manualmente con **Point** → **Add point at...** Otra alternativa es hacer un script abreviando los dos pasos anteriores:

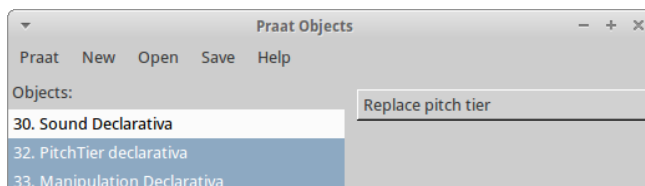
```
1 Create PitchTier: "declarativa", 0, 1
2 Add point: 0.080721, 129
3 Add point: 0.124198, 114
4 Add point: 0.273649, 124
5 Add point: 0.414042, 149
6 Add point: 0.548094, 110
7 Add point: 0.621461, 114
8 Add point: 0.764571, 114
9 Add point: 0.903153, 93
```

**Ventana 4.1.** Script 3

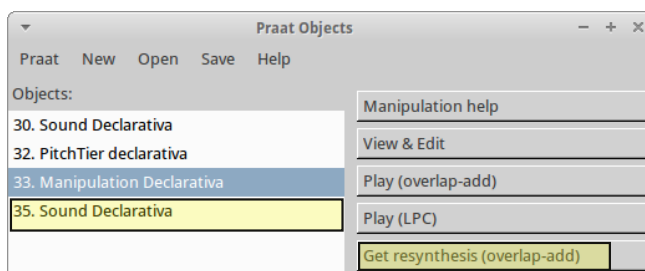
Ejecute el script (**Ctrl + R**) y obtendrá la curva estilizada:



- Quinto, debemos reemplazar la curva melódica por la curva melódica estilizada. Para ello es necesario crear un objeto **To manipulation** a partir del sonido original, y luego reemplazar la curva con el **PitchTier** que hemos creado en el punto anterior (**Replace pitch tier**):



- Finalmente, con **Get resynthesis (overlap-add)** convertimos el objeto **To manipulation** a un nuevo objeto **Sound** con el procesamiento de *resíntesis* de Praat.



El resultado final es una oración declarativa con una curva melódica *estilizada* y resintetizada. Si el procedimiento se realiza cuidadosamente, el sonido resultante será perceptivamente igual al original. Como se ve, la función **Manipulation** sirve para modificar habla natural, lo cual resulta de gran utilidad en estudios sobre la percepción del acento y la entonación.

## 4.7 MANIPULACIÓN DE LA DURACIÓN

En Praat también es posible modificar la duración de un sonido o un enunciado. Para hacerlo es necesario convertir el sonido original a un objeto **To manipulation** y abrirlo con **View & Edit**. El tiempo está representado por una línea roja que divide en el panel inferior del Editor. Para cambiar la duración de un sonido debe segmentarlo y ubicar los puntos que sean necesarios. Los movimientos ascendentes producen un incremento de la duración y los movimientos descendentes la reducen. Por ejemplo, en la figura que se muestra abajo manipulé dos vocales; la primera tiene la mitad de la duración y la segunda el doble. En este caso he usado cuatro puntos por sonido para no alterar el resto de la muestra.



El usuario dispone de las siguientes opciones en el Editor de manipulación:

1. **Add duration at point cursor:** añade un punto en el lugar en que está ubicado el cursor. Esta opción no permite controlar la magnitud del cambio, y sirve para añadir los puntos que delimitan el sonido.
2. **Add duration point at:** añade puntos con información temporal e información sobre la magnitud del cambio (*duración relativa*). Si desea reducir la duración a la mitad, la duración relativa debe tener un valor de 0.5; si desea incrementarla al doble su valor debe ser de 1.5.
3. **Remove duration point(s):** Elimina puntos temporales no deseados.

4. **Set duration range:** esta opción le permitirá cambiar el rango de manipulación. El valor estándar está entre 0.25 y 3.0, donde 1.0 es la duración normal.
5. **New duration:** introduce un nuevo panel y elimina el anterior.
6. **Forget duration:** elimina los puntos asignados.

#### 4.8 RECURSOS RECOMENDADOS

El lector puede acudir a otros manuales de Praat para complementar la información expuesta en este capítulo:

Kawahara, Shigeto. *Praat scripts*. Esta página contiene scripts para normalizar la amplitud, ajustar la duración y convertir archivos a mono. <http://user.keio.ac.jp/kawahara/resource.html#scripts>

Lennes, Mieta. *The Speech Corpus Toolkit for Praat*. En esta página la autora tiene a disposición **scripts** para remuestrear señales de audio, <http://www.helsinki.fi/~lennes/praat-scripts/>

Llisterri, Joaquim. *El análisis acústico del habla mediante Praat*. Tiene una sección en la cual aborda los temas presentados en este capítulo. Encontrará información adicional para manipular la intensidad. ENLACE: [http://liceu.uab.es/~joaquim/phonetics/fon\\_Praat/Praat.html](http://liceu.uab.es/~joaquim/phonetics/fon_Praat/Praat.html)

Román, Domingo. *Manual de Praat*. El capítulo octavo está dedicado a la *Manipulación del tono*. ENLACE: [http://domingo-roman.net/manual\\_analisis\\_acustico.html](http://domingo-roman.net/manual_analisis_acustico.html)

Yoon, Kyuchul. *Praat Tutorial*. Contiene scripts para manipular la duración y el tono en Praat.  
ENLACE: <http://www.ling.ohio-state.edu/kyoon/praat-tut/>



#### 4.9 EJERCICIOS SUGERIDOS

1. Tome un archivo de audio grabado con una tasa de muestreo de 44 100 Hz. Remuestree la señal a 8000 Hz y a 6000 Hz. ¿Cuál de las dos muestras transmite con más claridad la información? Argumente su respuesta en términos de los conceptos vistos en segundo capítulo.
2. Tome una oración interrogativa y una declarativa con la misma estructura gramatical, de tal manera que solo se diferencie una de otra por la entonación (puede usar las oraciones del capítulo anterior). Realice un etiquetado de los puntos de inflexión y estilice cada curva melódica. Durante el procedimiento de resíntesis intercambie la curva de la interrogativa por la curva de la declarativa y viceversa. Asegúrese que la duración de las oraciones sea el mismo y que concuerden con la duración del `PitchTier`.
3. Tome una oración o frase etiquetada segmentalmente (puede usar el fichero utilizado en el segundo ejercicio del capítulo anterior). Elimine el 75 % de la duración de las sílabas átonas o no acentuadas. ¿Puede afirmarse que hay una pérdida de la información lingüística? Investigue la literatura sobre el tema y explique los resultados.

**Cuadro 4.1.** Atajos de teclado para usar en el Editor VI

Opción en el menú	Teclas	Resultado
Add pitch point at cursor	Ctrl + T	Añadir punto en el lugar en que se ubica el cursor.
Remove pitch point(s)	Ctrl + Alt + T	Elimina uno o más puntos.
Stylize pitch (2st)	Ctrl + 2	Reduce la curva melódica a unos cuantos puntos de inflexión.
Interpolate quadratically	Ctrl + 4	La interpolación cuadrática unifica los puntos de inflexión como un contorno uniforme
Add duration point at cursor	Ctrl + D	Añadir punto en el lugar en que se ubica el cursor.
Remove duration point(s)	Ctrl + Alt + T	Elimina los puntos de duración.
Add pulse at cursor	Ctrl + P	Añadir una marca de pulso glotal.
Remove pulse(s)	Ctrl + Alt + P	Elimina uno o más pulsos.



## BIBLIOGRAFÍA

- Becerra, E. y Petersen, G. (2003). *Curso de lengua uitoto*. Universidad Nacional de Colombia, Facultad de Ciencias Humanas, Bogotá.
- Beckman, M., Díaz-Campos, M., McGory, J., y Morgan, T. (2002). Intonation across spanish, in the tones and break indices framework. *Probus*, 14:9–36.
- Boersma, P. (1998). *Functional Phonology*. Holland Academic Graphics, The Hague.
- Boersma, P. y Weenink, D. (2014). Praat: doing phonetics by computer [programa de computador]. versión 5.3.42. «<http://www.praat.org/>».
- Cantero, F. J. (2002). *Teoría y análisis de la entonación*. Edicions Universitat de Barcelona, Barcelona.
- Cutler, A. y Ladd, D. R. e. (1983). *Prosody: models and measurements*. Springer-Verlag, Berlín.
- Estebas, P., Eva y Prieto (2008). La notación prosódica en español. una revisión del Sp-ToBI. *Estudios de Fonética Experimental*, 18:263–283.
- Estruch, M., Garrido, J. M., Llisterri, J., y Riera, M. (2007). Técnicas y procedimientos para la representación de las curvas melódicas. *Revista de Lingüística Teórica y Aplicada*, 2(45):59–87.

- Face, P., Timothy y Prieto (2007). Rising accents in castilian spanish: a revision of Sp-ToBI. *Journal of Portuguese Linguistics*, 6 (1):117–146.
- Face, T. (2003). Intonation in spanish declaratives: differences between lab speech and spontaneous speech. *Catalan Journal of Linguistics*, 2:115–131.
- Fant, G. (1956). On the predictability of formant levels and spectrum envelopes from formant frequencies. En Halle, M., Lunt, H., y McLean, H., editores, *For Roman Jakobson*, páginas 109–120. Mouton, s-Gravenhage.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton, The Hague. Second Printing.
- Font, F. J., Dolors y Cantero (2009). Protocolo para el análisis melódico del habla. *Estudios de fonética experimental*, 18:17–32.
- Garrido, J. M. (1996). *Modelling Spanish Intonation for Text-to-Speech Applications*. Tesis Doctoral, Universitat Autònoma de Barcelona.
- Garrido, J. M. (2003). La escuela holandesa: el modelo IPO. En *Teorías de la entonación*. Ariel, Barcelona.
- Garrido, J. M., Llisterri, J., De la Mota, C., y Ríos, A. (1995). Estudio comparado de las características prosódicas de la oración simple en español en dos modalidades de lectura. *Phonetica. Trabajos de Fonética experimental*, 6:173–194.
- Harrington, J. (2012). Acoustic phonetics. En Hardcastle, W., Laver, J., y Gibbon, F., editores, *The Handbook of Phonetic Sciences*, Blackwell Handbooks in Linguistics, capítulo 3, páginas 81 –129. Wiley.

- IPA (1999). *Handbook of the International Phonetic Association*. Cambridge University Press, Cambridge.
- Johnson, K. (2006). *Acoustic and auditory Phonetics*. Blackwell Publishers, Cambridge.
- Klatt, D. y Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustic Society of America*, 87:820–857.
- Kohler, K. (1998). The disappearance of words in connected speech. *ZAS Working Papers in Linguistics*, 11:21–34.
- Ladd, R. (1996). *Intonational Phonology*. Cambridge University Press, Cambridge.
- Ladefoged, P. (1990). Some reflections on IPA. En *UCLA Working Papers in Phonetics*, 74, páginas 61–76. University of California, Los Angeles.
- Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. University of Chicago Press, Chicago.
- Lehiste, I. (1977). *Suprasegmentals*. MIT Press, Cambridge, MA.
- Liljencrants, J. y Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language*, 48(4):839–862.
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35(11):1773–1781.
- Llisterri, J., Joaquim y Mariño (1993). Spanish adaptation of SAMPA and automatic phonetic transcription. Informe técnico, ESPRIT PROJECT 6819 (SAM-A0). Speech Technology Assessment in Multilingual Applications.

- Martínez-Celdrán, E. (1991). *Fonética experimental: teoría y práctica*. Síntesis, Madrid.
- Moreno Cabrera, J. C. (1999). *Introducción a la Lingüística*. Síntesis, Madrid.
- Öhman, S. (2007). *The essence of Language: a philosophical problem: Why Noam Chomsky was never a linguist*. Nya Doxa, Nora.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Tesis Doctoral, Massachusetts Institute of Technology.
- Pluymaekers, M., Ernestus, M., y Baayen, R. (2005). Lexical frequency and acoustic reduction in spoken dutch. *Journal of the Acoustical Society of America*, 118(4):2561–2569.
- Port, A., Robert y Leary (2005). Against fomal phonology. *Language*, 81(4):927–964.
- Port, R. (2008). All is prosody: Phones and phonemes are the ghosts of letters. tomo Speech Prosody 2008, páginas 7–13. Campinas, Brasil.
- Prieto, P., Pilar y Roseano, editor (2010). *Transcription of Intonation of the Spanish Language*. Lincom Europa, Muenchen.
- Sosa, J. M. (1999). *La entonación del español. Su estructura fónica, variabilidad y dialectología*. Cátedra, Madrid.
- Stevens, K. (1998). *Acoustic Phonetics*. MIT Press, Cambridge, MA.
- Styler, W. (2012). *Using Praat for Linguistic Research*. University of Colorado at Boulder Phonetics lab, 1.1.1 edición.

- t ‘Hart, J., Collier, R., y Cohen, A. (1990). *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge University Press, Cambridge.
- Weenink, D. (2009). The Klattgrid acoustic speech synthesizer. En *Proceedings Interspeech*, páginas 2059–2062.
- Weenink, D. (2014). *Speech Signal Processing with Praat*. «<http://www.fon.hum.uva.nl/david/sspbook/sspbook.pdf>».
- Wells, J. (1997). SAMPA computer readable phonetic alphabet. En Gibbon, M. R., D. y Winski, R., editores, *Handbook of Standards and Resources for Spoken Language Systems*, 684-732, capítulo IV. Mouton de Gruyter, Berlin/New York.
- Wempe, T. y Boersma, P. (2003). The interactive design of an f0-related spectral analyser. En *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona.
- Wood, S. (1994-2014). Praat for beginners. <http://swphonetics.com/praat/introduction/>.





# ÍNDICE GENERAL

<b>Introducción</b>	<b>I</b>
<b>1 Fundamentos</b>	<b>1</b>
1.1 Interfaz gráfica . . . . .	1
1.2 Abrir y guardar archivos . . . . .	2
1.3 Grabación . . . . .	3
1.4 El objeto Sound . . . . .	5
1.5 Editor de sonido . . . . .	7
1.6 Cómo crear ondas sinusoidales . . . . .	9
1.7 Cómo crear ondas complejas . . . . .	11
1.8 Dibujando una onda . . . . .	15
1.9 Ejercicios sugeridos . . . . .	18
1.10 Bibliografía recomendada . . . . .	19
1.11 Atajos de teclado . . . . .	20
<b>2 Análisis del habla</b>	<b>25</b>
2.1 Espectrogramas . . . . .	25
2.2 Espectros . . . . .	28
2.3 Frecuencia fundamental ( $F_0$ ) . . . . .	32
2.4 Medición de $F_0$ . . . . .	35
2.5 Intensidad . . . . .	36
2.6 Medición de la intensidad . . . . .	38
2.7 Formantes ( $F_n$ ) . . . . .	38
2.8 Medición de $F_n$ . . . . .	41
2.9 To formant (burg...) . . . . .	43
2.10 Ancho de banda ( $B_n$ ) . . . . .	45
2.11 Medición de $B_n$ . . . . .	46
2.12 Creación de vocales con VowelEditor . . . . .	46

2.13 Pulsos glotales . . . . .	47
2.14 Duración . . . . .	49
2.15 Medición de la duración . . . . .	49
2.16 Dibujando un espectrograma . . . . .	51
2.17 Bibliografía recomendada . . . . .	54
2.18 Recursos recomendados . . . . .	56
2.19 Ejercicios sugeridos . . . . .	58
<b>3 Transcripción y etiquetado</b>	<b>63</b>
3.1 Segmentación . . . . .	63
3.2 El objeto <code>TextGrid</code> . . . . .	66
3.3 Búsquedas desde <code>Praat Objects</code> . . . . .	69
3.4 El Alfabeto Fonético Internacional . . . . .	71
3.5 Alfabeto SAMPA . . . . .	73
3.6 Transcripción y etiquetado de la entonación . . .	74
3.7 Etiquetado fonético de la curva melódica . . . .	75
3.8 Normalización de la curva melódica . . . . .	79
3.9 Transcripción fonológica . . . . .	82
3.10 Bibliografía recomendada . . . . .	86
3.11 Recursos recomendados . . . . .	87
3.12 Ejercicios sugeridos . . . . .	89
<b>4 Manipulación de la señal</b>	<b>93</b>
4.1 Edición de una señal digital . . . . .	93
4.2 Remuestreo . . . . .	94
4.3 Filtros . . . . .	95
4.4 To manipulation . . . . .	98
4.5 Manipulación del tono . . . . .	100
4.6 Estilización de la curva melódica . . . . .	101
4.7 Manipulación de la duración . . . . .	104
4.8 Recursos recomendados . . . . .	105
4.9 Ejercicios sugeridos . . . . .	106

# ÍNDICE DE FIGURAS

1.1	Ventana de objetos y ventana de imágenes . . . . .	2
1.2	Editor de sonido . . . . .	7
1.3	Onda sinusoidal . . . . .	9
1.4	Onda compleja . . . . .	12
2.1	Espectrograma de banda ancha . . . . .	26
2.2	Espectrograma de banda estrecha . . . . .	26
2.3	Ventanas de análisis espectral . . . . .	29
2.4	Espectro de banda estrecha, espectro de banda ancha y espectro continuo . . . . .	30
2.5	Visualización del contorno melódico . . . . .	33
2.6	Cálculo de $F_0$ . . . . .	36
2.7	Cálculo de $F_0$ en un espectrograma . . . . .	36
2.8	Curva de intensidad . . . . .	37
2.9	Error de cálculo de los formantes . . . . .	40
2.10	Espectros LPC de las vocales del español . . . . .	44
2.11	Ancho de banda y Nivel de intensidad . . . . .	45
2.12	Bandas de un espectrograma . . . . .	46
2.13	Pulsos glotales . . . . .	48
2.14	Vocales largas en la palabra /mo:ma/ ‘padre’ del uitoto ( <i>murui-bue</i> ) . . . . .	50
3.1	Segmentación de la interrogativa <i>y ¿pueden venir?</i>	64
3.2	Segmentación de la interrogativa <i>¿qué fue lo que te dieron?</i> . . . . .	64
3.3	Anotación con un <b>TextGrid</b> y cruce por cero . .	68
3.4	Unidades de análisis del sistema IPO . . . . .	76
3.5	Niveles de transcripción en el sistema IPO . . . .	78

3.6	Transcripción y etiquetado con el modelo de Análisis Melódico del Habla . . . . .	82
3.7	Transcripción y etiquetado En el sistema Sp_ToBI	85
4.1	Muestra en <i>mono</i> con un canal degradado . . . . .	93
4.2	Efecto de la opción Scale peak . . . . .	94
4.3	Filtros <i>paso-bajo</i> , <i>paso-alto</i> y <i>paso-banda</i> . . . . .	97
4.4	Editor de manipulación . . . . .	99

## ÍNDICE DE CUADROS

1.1	Atajos de teclado para usar con <b>Praat Objects</b>	21
1.2	Atajos de teclado para usar en el Editor I . . . .	22
1.3	Atajos de teclado para <b>Praat Picture</b> . . . . .	23
2.1	Atajos de teclado para usar en el Editor II . . . .	60
2.2	Atajos de teclado para usar en el Editor III . . .	61
3.1	Atajos de teclado para usar en el Editor IV . . .	91
3.2	Atajos de teclado para usar en el Editor V . . . .	92
4.1	Atajos de teclado para usar en el Editor VI . . .	107

PUBLICACIONES INSTITUTO CARO Y CUERVO  
SERIES MINOR

- XLIII *Manual de lingüística hispanoamericana. Tomo II: Notas para un seminario sobre el español americano. Coordinación académica y editorial de CÁNDIDO ARÁUS PUENTE.*
- XLIV JAIME BERNAL LEONGÓMEZ, *Panorama de lingüistas del siglo XX.* Presentación por Jaime Posada. Prólogo por GÉNOVEVA IRIARTE ESGUERRA.
- XLV MARIA STELLA GONZÁLEZ DE PÉREZ, *Manual de divulgación de lenguas indígenas de Colombia.* Prólogo de BELISARIO BETANCOUR.
- XLVI PIERRE ZIMA, *Para una sociología del texto literario.* Traducción del francés de CAMILO SARMIENTO JARAMILLO.
- XLVII *Lingüística hispánica, la gran riqueza del español: la variedad dentro de la unidad.* Coordinador de la obra JOSÉ MARÍA SANTOS ROVIRA.